DOCUMENT RESUME

ED 173 437	TM 009 576
TITLE	Proceedings of the Invitational Conference on Testing Problems (New York, New York, November 3, 1962).
PUB DATE	3 Nov 62
NOTE	129p.
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	*Cognitive Measurement; *College Entrance Examinations; Computer Oriented Programs; Computer Science; Creativity; *Creativity Tests; Educational Testing; *Equated Scores; Higher Education; Information Systems; Personality Assessment; *Psychological Testing; Tables (Data); *Testing Problems; Thought Processes

ABSTRACT

The first topic of the conference was creativity. Papers delivered were Knowledge and Creativity by Theodore C. Denise and Hobert W. Burns; Potentiality for Creativity and Its Measurement, by J. P. Guilford; and Some Methodological Issues in the Study of Creativity by Robert L. Thorndike. In the second session, a paper on comparability of scores was entitled: Can Useful General-Purpose Equivalency Tables be Prepared for Different College Admissions Tests? by William H. Angoff. The third session dealt with processing educational data and consisted of two papers: Information Systems in the Educational Environment, by Charles R. De Carlo; and The Role of Machines in Education Decision Making, by Ralph W. Tyler. The luncheon address, entitled Approaches to the Measurement of Intellectual Disposition, was delivered by T. R. McConnell. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED173437

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PROCEEDINGS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Educational
Testing Service
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM009 576

Copyright © 1963 by Educational Testing Service. All rights reserved.

Library of Congress Catalogue Number: 47-11220
Printed in the United States of America

Invitational Conference on Testing Problems

November 3, 1962
Hotel Roosevelt
New York City

ERIC F. GARDNER
Chairman

EDUCATIONAL TESTING SERVICE
Princeton, New Jersey
Los Angeles, California

ETS
Board of Trustees
1962-63

John D. Millett, *Chairman*

John S. Allen
Samuel M. Brownell
William G. Carr
John H. Fischer
John W. Gardner
Mason W. Gross
Stuart Hedden
Arnold E. Joyal
John I. Kirkpatrick
Katharine E. McBride
Edward S. Noyes
B. Alden Thresher
Robert J. Wert
Logan Wilson

ETS Officers

Henry Chauncey, *President*
William W. Turnbull, *Executive Vice President*
Henry S. Dyer, *Vice President for College Board Programs*
Robert L. Ebel, *Vice President for General Programs*
G. Dykeman Sterling,
    Vice President for Administration and Treasurer
Catherine G. Sharp, *Secretary*
John Graham, *Assistant Treasurer*

Foreword

THE 1962 Invitational Conference on Testing Problems explored a wide range of topics: theoretical discussions of creativity and how to measure it, the practical problem of equivalency tables for different admissions tests, and the technological question of the role of machines in education.

Speakers representing diverse interests and backgrounds discussed these subjects from the viewpoint of the philosopher, the psychologist, the educator, the research scholar, and the businessman. The result was an exchange of ideas and opinions that was stimulating not only to those of us in the field of educational testing but also to the many participants interested in other aspects of education who regularly attend this annual event.

For his work in planning and organizing the conference, we owe special thanks to the Chairman, Dr. Eric F. Gardner of Syracuse University. We are also grateful to Dr. T. R. McConnell of the University of California for his luncheon address on "Approaches to the Measurement of Intellectual Disposition" and to the other distinguished speakers whose papers appear in these PROCEEDINGS.

Henry Chauncey
President

Preface

DUE TO ITS REPUTATION as perhaps the outstanding meeting of the year for leaders in the testing field, the Invitational Conference on Testing Problems has attracted a steadily growing audience. The more than 700 leaders who attended the 1962 Conference represented the largest attendance to date. As the number of those attending has increased, so has the range of their special interests within the measurement field broadened. Both theorist and practitioner bring their varied needs and backgrounds to the Conference, and each anticipates a program that he will find personally rewarding. The present Chairman was very sensitive to the difficulty in providing a program appealing to such diversified interests.

Many of the past conferences have been organized about a central theme, and the various presentations have been devoted to different aspects of this theme. The present program has departed from this form of organization. Instead, several issues which I believe are important, current, and of varying degrees of technical complexity, were selected for consideration.

The first session of the Invitational Conference was devoted to the topic of intellectual creativity. Three speakers with differing interests and backgrounds were asked to address themselves to the following questions about creativity: What is it? How is it defined? How do we measure it? How is it related to intelligence? What have we learned about it through research?

Dr. Theodore C. Denise and Dr. Hobert Burns from Syracuse University, in their scholarly paper, "Knowledge and Creativity," considered some of these issues from the point of view of the professional philosopher. They provided an historical and philosophical background against which many of the current efforts in the measurement of creativity may be oriented and evaluated. Dr. J. P. Guilford presented some of the thoughts of a creative psychologist who has devoted many of his recent efforts to the problems inherent in the measurement of this elusive variable or, rather, variables. His paper, "Potentiality for Creativity and Its Measurement," reflected some of the work done by himself and his colleagues at the University of Southern California. The third speaker, Dr. Robert L. Thorndike, presented some very interesting re-analysis of data to make several cogent points in his paper, "Some Methodological Issues in the Study of Creativity."

The second topic was directed toward a very practical issue raised repeatedly by test users, in particular guidance counselors and college admissions officers. Their plea has been, "Since so many different tests are used, why doesn't someone provide conversion tables to convert a student's score on a particular test to comparable scores on others?" To the layman, this request sounds reasonable and appears to be one which can be granted easily. To the test theoretician, the request has numerous connotations and appears far from simple. In his paper, "Can Useful General-Purpose Equivalency Tables Be Prepared for Different College Admissions Tests?" Dr. William H. Angoff, of the Statistical Analysis Division of ETS presented a careful and scholarly analysis of the technical problems involved in determining equivalency tables.

At the luncheon meeting, Dr. Henry Chauncey presented Dr. T. R. McConnell, Director of the Center for the Study of Higher Education, University of California (Berkeley), whose description of some of the important research findings obtained by himself and his staff was most interesting.

The afternoon session turned to the theme of "Processing Educational Data." The speakers were Dr. Charles R. DeCarlo, Director of Education, International Business Machines Corporation, who spoke on "Information Systems in the Educational Environment," and Dr. Ralph W. Tyler, Director, Center for Advanced Study in the Behavioral Sciences, who spoke on "The Role of Machines in Educational Decision-Making." Both papers, which were interesting and provocative, served as the basis for considerable participation by the audience.

I wish to express my sincere appreciation and thanks to the speakers for the considerable amount of time and thought involved in the preparation of their remarks, and for their excellent presentations.

It is my pleasure also, on behalf of the profession at large and the members of the Conference, to thank the Educational Testing Service for its sponsorship of this annual event, which represents such an important forum for the exchange of information and ideas about problems in educational measurement.

Eric F. Gardner
Chairman

CONTENTS

iii Foreword by **Henry Chauncey**

v Preface by **Eric F. Gardner**

SESSION I: Creativity

13 Knowledge and Creativity, **Theodore C. Denise**, **Hobert W. Burns**, Department of Philosophy, Syracuse University

31 Potentiality for Creativity and Its Measurement, **J. P. Guilford**, Department of Psychology, University of Southern California

40 Some Methodological Issues in the Study of Creativity, **Robert L. Thorndike**, Teachers College, Columbia University

SESSION II: Comparability of Scores

57 Can Useful General-Purpose Equivalency Tables Be Prepared for Different College Admissions Tests? **William H. Angoff**, Statistical Analysis Division, Educational Testing Service

CONTENTS continued

LUNCHEON ADDRESS

74 Approaches to the Measurement of Intellectual Disposition, T. R. McCONNELL, Center for the Study of Higher Education, University of California (Berkeley)

SESSION III: Processing Educational Data

91 Information Systems in the Educational Environment, CHARLES R. DECARLO, International Business Machines Corporation

102 The Role of Machines in Educational Decision-Making, RALPH W. TYLER, Center for Advanced Study in the Behavioral Sciences

114 List of Conference Participants

Session I

Theme:
Creativity

Knowledge and Creativity

THEODORE C. DENISE and
ROBERT W. BURNS,
Department of Philosophy,
Syracuse University

A SURVEY of the literature on creativity makes clear that two major problems in the investigation of creativity have yet to be solved: the criterial problem of determining and defining the nature of creativity, and the methodological problem of evaluating and measuring creativity.

It is equally clear that academic competence in philosophy does not in itself yield solutions to these problems, for such solutions lie more in the intellectual jurisdiction of psychology than of philosophy. Still, even a Spartan mentality could perceive that the intellectual boundaries of philosophy and psychology run through some common territory and, consequently, that our areas of inquiry are sometimes adjacent. One such area involves the theory of knowledge, which frequently appears in psychological literature under the rubric of "cognition," while philosophers use the term "epistemology." It is from this common touchstone, found lying in the field of philosophy of science, that we have based the following comments on creativity.

Our thesis is simple: contrary to some literature which frees creativity from the *terra firma* of past experience and seems to base it on a metaphysic of ideational free-wheeling and transempirical dealing, we believe that creativity is cognitively based. We suggest that creativity cannot be separated from knowledge, nor the creative process from the knowing process, and therefore that a thorough analysis of creativity must open and close with epistemological considerations.

In the pages to follow we propose to elaborate on this thesis by reviewing, briefly, the philosophic history of creativity as seen by the epistemologist. This is best done, given the time allotted us, by contrasting two major streams of epistemological thought – rationalism and empiricism – and concluding with a comment we hope will be of interest to psychological theorists.

I

René Descartes, the man who first realized that every algebraic equation may be represented geometrically and who first recognized the full philosophical purchase of *Cogito ergo sum*, comments thusly in his *A Discourse on Method:*

"For myself, I have never fancied my mind to be in any respect more perfect than those of the generality; on the contrary, I have often wished that I were equal to some others in promptitude of thought, or in clearness and distinctness of imagination, or in fulness and readiness of memory. And besides these, I know of no other qualities that contribute to the perfection of the mind; for as to the reason or sense, inasmuch as it is that alone which constitutes us men, and distinguishes us from the brutes, I am disposed to believe that it is to be found complete in each individual; and on this point to adopt the common opinion of philosophers, who say that the difference of greater and less holds only among the accidents, and not among the *forms* or *natures* of *individuals* of the same *species.*"

How pleasant the thought: Descartes, the universal genius, places us at his own level with respect to reason! But how short the pleasure: He has made every fool our peer. Is Descartes to be taken seriously in this paragraph? Do not the facts of his own life refute his words? We know, for example, that as a young student in the Royal College of La Fleche he was granted the privilege of remaining in bed until a late hour in the morning in order to think and reflect. It was during one such morning that the idea of analytical geometry flashed through his mind. Is it too much to conjecture that Descartes' Jesuit mentors felt that he had given evidence of an ability for profound and original thought? That he was, as we say, an identifiably "gifted" and "creative" young man?

The thrust of Descartes' epistemology, the thrust of classical rationalism, is not, however, parried by a foil so lightly grasped. We have quoted from remarks prelusive to the rationalistic argument, first, that *genuine* knowledge takes the form of a system of statements derived deductively from indubitable first principles, and, second, that these incorrigible truths lie equally available for every man, their discovery depending solely upon his reason. In classical rationalism the apprehension of truth is absolutely irreducible in either whole or part to keenness of the senses, to verbal facility, to pictorial inventiveness, or to excellence of memory. Our senses may occasion thoughts, our words may express them, our imagination may illustrate them, and our memory may represent their logical relationships sequentially; but, perceiving is *not* understanding, speaking is *not* thinking, imagining is *not* conceiving, and remembering is *not* recognizing logical relationships.

But what of error and ignorance? If we grant that truth is a purely conceptual category which cannot be approached through the senses or through those faculties of mind which depend on them, it would seem either that all men must in fact have equal knowledge or that they must differ with respect to reason itself. The rationalists escape through the horns of this dilemma by insisting, in the first place, that all men are, in various ways, diverted away from exclusive reliance on reason, and, in the second place, that an individual's knowledge varies inversely with the extent to which he is thus diverted. It is the precise concern of Descartes and Spinoza in their methodological treatises to show us how these diversions may be precluded, that is, to show us how we may free ourselves from irrational influences. For them the terms "method of discovery," "improving the mind," and "freedom of thought" all pertain to achievements in sustaining rationality and coming into possession thereby of that modicum of genuine knowledge which is possible for human beings. The terms of classical rationalism do not imply individual differences in our capacities for true knowledge, nor that *knowledge is in any respect a product of human inventiveness or creativity.*

Before ranging too far beyond Descartes' comment that "the difference of greater and less holds only among the accidents, and not among the *forms* or *natures* of *individuals* of the same *species*," it is interesting to note that this "common opinion of philosophers" inherited from Plato and Aristotle still *serves today to set creativity apart from knowledge.* For example, in his well-known essay, "The Dehumanization of Arts," an apologetic for contemporary art, José Ortega y Gassett writes:

"If the new art is not accessible to every man this implies that its impulses are not of a generically human kind. It is an art not for men in general but for a special class of men who may not be better but who evidently are different."

It is precisely an accidental quality or ability, the ability "to construct something that is not a copy of 'nature' and yet possesses substance of its own" which defines artistic genius. Notice that insofar as a rationalistic epistemology lingers — as it indeed does in the philosophy of Ortega — the claim being pressed here must be for *artistic* genius and not for *scientific* genius. The ability to alter nature, to reconstitute it through creative acts of imagination, lies over and beyond the ability to know it, the ability to understand it.

II

But what of rationalism's great alternative, empiricism? Is the tradition of empiricism rooted in the belief that creative acts of intelligence are integral to the knowledge equation? The answer is that classical empiricism repudiates this belief as surely as does classical rationalism. Both Descartes' older and younger contemporaries, Francis Bacon and John Locke, assumed that the mind at birth is a *tabula rasa,* a preconceptionless instrument recording sense impressions. In his *The Great Instauration,* Bacon goes so far as to suggest by certain of his metaphors that in freeing the mind from those predilections which interfere with and distort its reception of data from the senses, we achieve not only the *necessary* condition for knowledge but also the *sufficient* condition for knowledge as well. Creativity is not only excluded but precluded from Bacon's epistemological equation. He says that a mind is "like a mirror with a true and even surface fit to reflect the genuine way of things," that civilized man is typically so beset by the mechanisms of false opinion "that there is no true and even surface left to reflect the genuine rays of things," and that we must first restore this surface by purging ourselves of such mechanisms and then "make provisions for finding the genuine light of nature . . . and bring it to bear."

The connection between the *activity* of removing corrupting influences from the mind and gaining thereby that *passive* state of mind to which knowledge is alone natural has been affirmed again and again in varying contexts throughout the history of thought. Its greatest and most influential expression is in Plato and its most recent major expression is in modern romanticism, but it is also expressed in the early Renaissance, in the Reformation, in classical rationalism and, as we are discussing, in classical empiricism. At one level it is instructive to state rationalism and empiricism in terms of this formula: The rationalist is saying that all attempts to found knowledge on sense-experience stand between us and genuine knowledge; the empiricist is saying quite the reverse, that all attempts to found knowledge on *a priori* ideas preclude genuine knowledge. Each is saying of the other that he wrongly conceives the object of knowledge, that he mistakes his own speculations for discoveries of truth, that he allows his ability to *create* to take precedence over his ability to *reflect*.

At another level, however, it would seem that the thesis of a natural state of knowing cannot be equally embraced by rationalists and empiricists. After all, the classical empiricists must somehow account for our moving from particular sense-experience to those general principles we depend upon in our science. He does not claim for the latter that we discover them fully developed and self-attesting. He must account for an *inductive* movement of mind.

The history of classical empiricism in this regard is complex and cannot be analyzed here; it will suffice to mention that after Bacon the problems of induction led empiricists to renounce certainty as the mark of genuine knowledge. This fact does not indicate that there was a shift to support the thesis that creativity is linked with knowledge, however. From the elementary Bacon to the sophisticated Hume and the neo-Baconian Mill there was a continuing effort to reduce knowledge to little more than a set of mental routines.

The Baconian rules for discovery are as meticulously worked out as are the Cartesian rules and they foreclose against human creativity almost as completely. Furthermore, in following the Aristotelian doctrine of essential characteristics, that is, in following the doctrine that each kind of thing has a basic quality which, when discovered, is scientifically explicative, Bacon represents science as a terminal set of discoveries rather than as an open-ended enterprise.

III

Before shifting attention to two sequences of events which are commonly believed to show both the classical rationalists and the classical empiricists to be holding inadequate views of science, let us review the connection between creativity and knowledge as it was generally conceived to be during the Enlightenment, that eighteenth-century period when empiricism and rationalism were blended together with blissful unconcern for the philosophical difficulties involved.

The assumptions of the eighteenth century were that rational laws govern the universe in both its natural and human domains and that each man has sufficient rational capacity to discover, understand, and abide by them as they are revealed in experience. Technically speaking, the term "rational capacity" in this context is nearer to the empiricists' term "common sense" than it is to the rationalists' term "possession of *a priori* knowledge"; but, popularly speaking, the suggested synonym "common sense" takes on the rationalistic claim — and indeed extends it — that man has the capability of discovering and knowing that he has discovered the verities of the universe. "Use your common sense" commands one to look without prejudice and with the courage to see, and assures him that by so looking he will see and that by so seeing he will understand. Through the use of common sense, each of us is capable of really seeing an apple fall just as Newton did, that is, each of us is capable of perceiving the events of this world as intelligible rather than as random happenings.

A twentieth-century commentator, Lawrence Binyon, remarks as follows about the eighteenth century:

"No period had more the courage of its taste. Genius, however overpowering, was not to have its excesses and licenses condoned. Neither Shakespeare nor Milton intimidated Dr. Johnson: each must submit to the tribunal of reason."

We have only to consult the philosopher Diderot to find the Enlightenment's explanation:

"The poet who feigns, and the philosopher who reasons, are equally, and in the same sense, logical or illogical, for to be logical and to be aware of the necessary chain of phenomena is the same thing. This suffices, it seems to me, to show the analogy between fiction and truth, and to characterize poet and philosopher."

Note that man's essential feature, his rational capacity, is now stretched to underlie not only his efforts in science and philosophy but also his endeavors in the arts as well. *A creative act of imagination must at the least conform to a possible pattern of reality to have esthetic worth.* In the most extreme moments of the Enlightenment, this thesis is taken to assert that experience plus reason, that is, common sense, reveals the laws of taste just as surely as it reveals the laws of mechanics; that in the final analysis a candidate work of art – its claim for originality being beside the point – must be judged genuine or bogus in terms of the appropriate laws just as a purported physical happening must be judged possible or impossible by an appeal to the laws of mechanics.

The gulf between art and knowledge, so lamented in our times, was eliminated in the Enlightenment by subsuming the former to the latter; but to think that this unification was achieved by finding an essential role for creative thought in the human endeavor to know, is a mistake. Creativity was not added to science; rather, it was largely subtracted from art.

IV

"There is . . . but one matter in the whole universe, and we know this by the simple fact of its being extended." With these words, Descartes declared against the distinction between space and bodies-in-space as essential. Insofar as there are space and body differences and differences between bodies with respect to hardness, weight, and the like, these differences depend upon motion. However strange Descartes' famous denial of nonmaterial space may sound to us, it constituted a brilliant move for him to make. He was working within a heritage in which – strict Platonists excepted – geometry was understood to be the study of the extensional characteristic of bodies and in which the problem of how geometry was related to nonmaterial space was scarcely posed. Descartes accepted the prevailing view, saw the problem, and obviated it by denying the distinction in kind between bodies and space. This denial of a prevailing distinction also assisted him in his metaphysical dualism by removing the confusing conception of a space which is at once nonmental and nonmaterial. With respect to physics, his declaration that space and body are of a kind enabled him to initiate a mathematical theory of mechanics which did not involve the intuitively repugnant idea of "action at a distance." Finally, by subsuming space to matter, Descartes extended the knowledge claim for Euclidean geometry, the *a priori* science *par excellence.*

With Euclid's system mathematically united with algebra and philosophically extended to underlie the impressively general and precise new mechanics, with the self-evidence of Euclid's axioms and postulates generally conceded, with tradition supporting the notion that Euclid's theorems pertain to the spatial properties of bodies, with no alternative geometries extant, and with Descartes and Spinoza scoring notable successes with the *a priori* method in metaphysics proper, the case for rationalism was at its peak. Who was to challenge its thesis that knowledge of reality is innate knowledge? That to know anything genuinely is to experience a disclosure of reason without recourse to sense-experience: That the capability for knowing as distinguished from actually knowing is the same in all men?

History's answer to this question is Gerolamo Saccheri, an Italian mathematician who, early in the eighteenth century, attempted a *reductio ad absurdum* proof of Euclid's now famous parallel postulate. The self-evidence of this postulate – but not its truth – has been questioned occasionally since at least the fifth century; this had led to a variety of efforts to show the status of the postulate within the system, whether as postulate or as theorem. Saccheri's approach to the problem was to replace the postulate with contrary ones and then to deduce contradictions; succeeding in this he would have shown the true status of the postulate to be that of a theorem. What he in fact did deduce were psychologically strange propositions rather than logically contradictory ones. Unfortunately for his immortality, he did not see the mathematical import of what he had done and failing this he could not see its philosophical import, that is, he did not see that he had produced the beginnings of non-Euclidean geometries and therefore he could not see that the tests of reason without reference to sense-experience were no longer decisive with respect to the nature of reality. The machinery for the destruction of rationalism in its extreme or classical form was on hand, however, and its demise became merely a matter of time.

Let us take a closer look at the critique of classical rationalism occasioned by Saccheri; it will serve our search for a connection between creativity and knowledge. Grant to begin with that all of Euclid's primitive propositions, with the exception of his parallel postulate, are indeed self-evident. The fact remains that the latter is but one among a set of possible contrary propositions none of which may be deduced from the self-evident primitives each of which is consistent with them, and the absence of all of which renders the self-evident primitives inadequate as a physical geometry. Thus Descartes' hope that complete knowledge of reality can be obtained from a few incorrigible principles is vitiated. His tests of reason — self-evidence, deducibility, and logical consistency — are shown to be inadequate in and of themselves for knowing reality.

But is not our general confidence in the possibility of self-evident truths which pertain directly to reality also severely taxed? After all, the parallel postulate had been conceded by most to be self-evident and from an everyday point of view it is still so judged. If one "self-evident" truth can give way, why not another? What *is* the test of self-evidence? If it is that any proposition denying a given proposition is either unimaginable or inconceivable, then the test seems to be psychological and to depend on our having self-awarely exhausted human ingenuity. But how can we know when we have done so? Is the failure of one group of men to conceive the opposite of a proposition *necessarily* the sign that a subsequent group will fail? The history of thought records ever-enlarging conceptual contexts – witness the reluctant Saccheri – and, as a result, ever-enlarged possibilities for denying received self-evident truths.

If the test of self-evidence is instead that the denial of a given proposition is impossible in the sense that the denial is self-contradictory, then clearly the test is not psychologically contingent. There are complications even here about which contemporary philosophers are concerned, but, at the very best, the rationalist can claim only that logical truths – analytical truths – are self-evident. However, such truths lack content. They do not refer; and since they do not, they make no direct claim upon reality. They may indeed hold for the formal languages which we require if we are to have any knowledge whatsoever about reality, but that is quite another matter.

These criticisms, and others which might be adduced, show self-evidence to be both suspect and insufficient as the test of knowledge, even when bolstered by the purely logical tests of validity and consistency. Furthermore, with self-evidence discredited, the classical rationalist loses his basis for depicting reason as that passive and unique source of truth which is equally and automatically available to all men who free themselves from irrational distortions. If, in the spirit of rationalism, we retrench by claiming that reason is entirely reflected by the purely logical tests, validity and consistency, which remain, then we restrict reason to the domain of the *form* of knowledge as opposed to the *content* of knowledge. In so doing we narrow the term beyond the warrant of convention, and, in any event, we are forced to supplement it with new terms to describe our actually thinking about things. If, on the other hand, we retain the term "reason" as totally associated with knowing, it must be enlarged to accommodate our acts of proposing content as well as our acts of determining that such proposals are logically possible.

Thus a critical review of classical rationalism forces us to admit that the knowledge equation involves our attempting accounts of reality in psychological contexts of not knowing, or, stated more precisely, in psychological contexts in which we seek to know without the availability of *a priori* ideas. Granted that the results of our seeking to know are to be judged as irrational when they are shown to be illogical, it remains that man's being rational – in the sense at least of his having the capacity for a modicum of knowledge about the world – *presupposes a kind of human inventiveness or creativity which the classical rationalists condemned as irrational.*

V

But what of the other great epistemological tradition we have been examining? Must it too be brought to terms with creativity? A sequence of events under way, even as Bacon wrote, makes it clear that such an accommodation is called for.

Many of the successes of Renaissance thinkers came about through first reviving ancient ideas and then perfecting them. Copernicus affords a brilliant example: He rediscovered Aristarchus' heliocentric theory of astronomy and gave it successful mathematical expression. The historic conflict between religion and science that this event set off is well-known and requires no direct comment, but there are connected features which do.

The accepted Endoxus-Aristotle-Ptolemy geocentric theory had come to be regarded as an actual or real account of celestial happenings. This view was justified by attending to the position of astronomy in Aristotle's metaphysics as a theoretical or certain science rather than as a practical or approximate science, by neglecting Aristotle's tentativeness with respect to the details of the science, and by belief in revelatory celestial references. Now the earlier astronomers themselves had viewed their accomplishments, if not their science, quite the other way round. Even Ptolemy looked at his own account as doing no more than filling the practical needs of predicting and explaining observables. For him, Copernicus' account would have been merely a simpler and possibly more accurate *as if* account of the movement of the planets. The astronomer is satisfied to develop serviceable fictions; he lacks both the motive and the method to know the heavens as they really are.

Copernicus revived the astronomers' effort without reviving their philosophical modesty and thereby confronted himself, his student Rheticus, and later, Kepler and Galileo with a philosophical puzzle of the first order, namely the problem of justifying the new astronomy as real. All of them except Galileo in the final period of his thought accepted the dictum from the past that genuine knowledge is certain knowledge rather than approximate knowledge. But how can certainty be claimed for a scientific theory when the grounds for accepting it are no more than its logical propriety and its confirmation by observation? Both Copernicus and Kepler sought to bolster their position by instituting another test for scientific truth, the test of divine simplicity. God is *the* great mathematician and He would be less than such if He had the planets move in any but the most mathematically simple and general way. But this test is itself either question-begging or metaphysically presumptive in the extreme.

In the final analysis, the Renaissance realists were faced with three alternatives: (1) The method of science, the hypothetical-deductive method, might be reconceived in such a way that sense-experience could somehow assure certitude; or (2) the method might be shown supplemental to another method, a method of certainty; or (3) the requirement that only a method of certainty can give knowledge of reality might be set aside. Bacon chose the first alternative, Descartes the second, and Galileo the third. History has declared for Galileo's choice, in part, at least, because neither Descartes nor Bacon was able to make good on his. As we have seen, Descartes' difficulties could scarcely have been anticipated; this was not so with respect to Bacon.

Aware of the objections on logical grounds to claiming certainty for a theory simply because its consequences are affirmed, Bacon sought to draw instead upon the impeccable argument-form that the denial of a theory's consequences constitutes a denial of that theory. But a method which involves an inference from the falsification of all theories but one in a group to the truth of that one, requires extraordinary assumptions of accommodation. One of Bacon's assumptions is that only a limited number of theories provide possible accounts of a given phenomenon. However, granting the formal situation Bacon envisions, it remains that he is forced to make an assumption concerning the connection between theories and reality in order to apply his method to elimination. In making this assumption Bacon unwittingly misses the basic scientific insight of the Renaissance – namely, that *the central feature of a scientific theory depends upon dealing with a quality, a property of a thing, as a quantity;* of, for example, converting the property of moving into a rate of movement or the property of being hot into a degree of temperature. Bacon's assumption is instead that the central feature of a scientific theory resides in identifying the essential or defining quality of things of a kind from among those qualities we observe them to have. With meticulous observation of numerous instances under different conditions we may eliminate all erroneous generalizations in favor of that generalization that is suitable.

All additional assumptions of accommodation in force – for example, the asssumptions that our senses can be veridical and that all relevant qualities of objects are represented in sense-experience – Baconian science can have no more power to explain and predict than a set of real definitions. Bacon loses touch with science as we know it precisely because of his quest for certainty; understandably impressed with the irrefutability of immediate sense-experiences, he attempted to portray them as the sole building blocks of scientific theories. But the classical empiricists' hope to avoid drawing capital from human inventiveness with all its fallibility proved as short-lived as the classical rationalists'. Bacon could not eliminate the human variable of creativity from the epistemological equation.

VI

Our criticisms of classical empiricism and classical rationalism must not lead us to neglect their wisdom; purged of its claim for certainty, each epistemology captures an essential feature of scientific inquiry. Bacon and his followers saw clearly the need for systematically selecting and testing explicative qualities from among groups of observable qualities. What they failed to see was that fruitful explanation depends upon dealing with a quality as quantity rather than as quality. Quantitative data gained on successive occasions and under varying conditions fix a scientific problem as one of somehow accounting for these data and anticipating future data. The secret of "somehow account for . . . and anticipating" involves, in the first place, an abstracting process in which expressions of empirical measure are regarded as mathematical expressions occurring within a logico-mathematical structure. In the second place, it involves an effort to suggest more general mathematical expressions permissible within that structure from which these genetically empirical expressions may be deduced along with still other expressions destined for empirical testing.

The Cartesians saw the need for converting qualities to quantities and moving thereby to gain the deductive power of mathematics. But their preoccupation with self-evidence blinded them to all but this crucial middle movement of the scientific method as we conceive it to be: they did not, as the empiricists did, see the initial movement of data-gathering and the final movement of confirming. In this sense then, the sense in which modern scientific method is analyzable into the conjunction of the three movements mentioned, it is a synthesis of classical empiricism and classical rationalism, the issue of their basic approaches to knowledge. *It is an issue, however, which, unlike its parents, lays stress upon human creativity at every point.*

Galileo not only became the exemplar of the new method through his work, but he also gave it preliminary expression, and, perhaps, its ultimate defense. Philosophically speaking, his claims represent at once a retreat from and an advance beyond scientific realism in its early form, and his claims, even as his example, remain central in the thinking of most scientists today. His retreat amounts to saying that a scientific theory can never provide us with more than sufficient reasons for the occurrence of a phenomenon or the propriety of a law. That is, a theory is not a statement of necessary and sufficient reasons, *the* reasons, either for anything which really happens or for a law which really prevails. This means that science is, as a matter of principle, open-ended, that a "final" theory is merely one which seemingly lies beyond our inventiveness to supplant.

VII

The Galilean thesis is, of course, too pat to enjoy philosophical invulnerability. The history of thought illustrates this impressively. But, philosophical criticism notwithstanding, Galileo's general claim for science as knowledge survives, protected as it is by a wall of practical successes grown to inestimable height. The point which follows from these observations for our present attention is this: *As scientific method was instated to primacy in knowing, human creativity was itself elevated to an essential role.*

If, however, as we have labored to show, scientific method depends on creativity, *it is equally the case that creativity, insofar as it enters into knowledge, does so by way of and under the control of scientific method.* There is a sense in which to speak of controlling creativity constitutes a contradiction in terms, but there is another sense in which this is not so, a sense in which we mean that *creativity is invited into play at prescribed stages within a disciplined quest for knowledge and its product judged by the objective standards of reason and sense-experience.*

It was the practice of the distinguished scientists of the immediate past to speak of theories as "free creations of the mind" or as the "products of human genius." As a disclaimer of the confining accounts of classical rationalism and classical empiricism such expressions are clearly warranted. But they can be misleading where there is no need to disclaim because they neglect the contextual realities of acts of scientific genius. Drawing a positive line between totally nondirective innovation — an extreme of freedom which many of today's artists are demanding for themselves — and context-bound innovation, which stands at a single remove from such acts of intelligence as deductive inference, enthymematic discovery, and inductive generalization, we may say: that a scientific theory is an invention, a creative act, the basic terms of which are not necessarily established by an experimental situation; and the validation of which may be achieved by indirect confirmation in sense-experience when direct confirmation is on-principle or in-practice impossible.

Let us add token illustrations to show both how thoroughly scientific knowledge is permeated with human creation and how the creative acts of science represent free responses to highly structured situations within a settled method rather than unrestricted flights of admirable fancy. While attending a service at the Cathedral at Pisa, Galileo, then a medical student, watched a suspended lamp, pulled to one side for lighting, swing back and forth in ever-diminishing arcs. He timed these with his pulse and was in a position to formulate his famous empirical law of the pendulum, a law later shown explicable by still more general laws. The creative act of converting a quality into a quantity is decisive here, and it is an act of a kind called for again and again in the development of science. It is called for most dramatically by such sciences as psychology and the social sciences which sit conspicuously at the edge of qualities for which the claim of nonconvertibility is strongest. Color is a case in point: it could not be completely measured until Horrower and Koffka were able to conceive of and demonstrate a third dimension, saturation, to be measured along with hue and brilliance. That the movement of data-gathering is not always mundane, that it may become possible only after a major creative act, cannot be clearer than it is to those who desire to place creativity itself under scientific scrutiny.

The role of creativity is most obvious in the middle movement of scientific inquiry. In this connection we have but to mention the name Einstein. When faced with the seeming facts that an observer cannot detect his own motion and that the velocity of light is independent of the motion of its source, he reasoned that the Newtonian postulate in which measurements of mass, length, and time are the same for all observers, stands in contradiction. His theoretical problem was clarified by logical and mathematical analysis; he labored at logical and mathematical reconception until he hit upon his celebrated Special Theory of Relativity in which mass, length, and time measurements are taken to be relative to each observer. It is worth insisting again that even a revolutionary act of creation in science, such as we know Einstein's to be, does not proceed with the amalgam of reasoning and experimentation by which the difficulty to be overcome is isolated and made precise, nor without command of the received tools for resolution. It is true for Einstein — no matter his creativity quotient — as it is for the rest of mankind that creation *ex nihilo* falls to God alone.

VIII

Our major conclusion, then, is that creation does not occur in an intellectual vacuum but requires a firm epistemological base.

Neither rationalism nor classical empiricism provides an adequate basis, for neither reason alone nor unmediated experience yields knowledge. Thus the cognitive base of creativity was lacking until the twin-birth of modern science and empirical philosophy. With the growth and development of these intellectual twins, each characterized by quantification, the epistemological preconditions for an analysis and investigation of creativity are present.

We also promised a comment pertinent to psychological theory, and it is this: if inquiry into creativity is to be deliberately guided by a psychological theory, and if the major theories are psychological field theorism and modern behaviorism, then we do not see how our epistemological considerations would lead to, or imply, field theorism. This is not to say that our philosophic considerations clearly imply a behavioristic approach; it is only to note that they do not rule out modern psychological behaviorism as a theoretical tool to be investigated for its efficacy as a guide in seeking solutions to the criterial and methodological problems of creativity.

While time inhibits a full disclosure of the reasoning which prompts our suspicions in this respect, a sense of responsibility suggests that we should at least outline our thought for your analysis.

First, as argued in our text, the central feature of scientific theory depends upon dealing with a quality, a property of a thing, as a quantity; thus quantification is the *sine qua non* of empirical science, and insofar as psychology is scientific it relies on quantification.

Second, it does seem that a science based on heuristic logico-mathematic techniques will likely be more productive than one which is not; thus psychology, as the most advanced behavioral science, would rely on the more fruitful statistical tools.

Third, if, as philosophers, we take the word of such eminent psychological theorists as Estes, Hull, Sears, Stevens, or Tolman as valid, then field or phenomenological psychology dispenses with S-variables in theory, with the result that the laws achieved are of the R-R type; but since modern behaviorists tend to emphasize the objective, environmental conditions that are represented by the S-variable, the laws they achieve are of the S-o-R variety.

And fourth, as Bergman and Spence report, it is logically and mathematically possible to deduce an R-R correlation from a set of S-R correlations while, conversely, it is not possible to deduce an S-R correlation from a set of R-R correlations. This would suggest that S-R behaviorism might well have a significant quantitative advantage.

On the basis of such considerations about methodological procedures and possibilities it may be that an S-R theoretic model has more fruitful explanatory power than does an R-R model. Even if this is the case, it is not to say that phenomenological description is of no value to an investigation of creativity; it may be of great value, even if it cannot generate causal explanation. But if we are to seek out causal explanations in the search for solutions to the problem of creativity, modern behavioral theory should not be ruled out.

We suggest that behaviorism not be discarded on *a priori* grounds in the analysis of creativity, for it appears that it, rather than field theorism, is on firmer epistemological grounds.

Potentiality for Creativity and Its Measurement

J. P. GUILFORD,
*Department of Psychology,
University of Southern
California*

"CREATIVITY," like "love," is a many-splendored thing. Small wonder that few have ventured to define it. At a conference on creativity a few months ago, each of the thirty-odd members was asked to write a list of his free associations to the word "creativity." The results were almost as varied as the personalities of those present.

We gain a little bit if we take a first step toward discrimination, namely, the distinction between creative potential and creative production. Creative production, in the popular sense, is the aspect that catches general public fancy because the creative person's output is so frequently in the form of tangible products, such as a poem, a novel, a musical composition, an invention, a painting, a scientific theory, or a philosophical system.

To a psychologist who is interested in the mental operations that lead up to the emergence of the creator's product, the criteria of creative production must be of a different kind. There is agreement that the tangible product must have some novel aspects. The thinking that leads to that product also has novel aspects. In either case, it is these novel aspects that justify the label "creative." We are not concerned here with the operational criteria by which novelty can be gauged. What I wish to point out is that on the way to his final, public product, the creative thinker arrives at numerous *psychological* products. In focusing attention on the public product, we overlook the numerous ideas that the inventor had and discarded. From the psychological point of view, those generated ideas also have many chances of being novel.

In simplest terms, an individual's potential for being creative is his readiness to produce novel ideas or psychological products. In this we should include the production of old ideas in new connections. His readiness depends upon many things. An essential part of his preparation is in the form of specific items of information that are available to him from his memory storage. But, as we know, having the information is not sufficient. It is what the person does with the information that is important. When we speak of creative potential, we usually have in mind the dispositions that enable a person to use his information in new ways. Some of these dispositions are abilities or aptitudes while others are traits of interest, needs, attitudes, and of temperament. The emphasis upon aptitudes in this paper should not be interpreted to mean that these other kinds of traits are thought to be of negligible consequence.

Our investigations of creative potential at the University of Southern California have been directed mainly at the aptitude aspects, in part because they have been conducted in an Aptitudes Project and in part because of the belief that an understanding of the aptitudes will also give us much information regarding the nature of creative thinking itself. From the standpoint of aptitudes alone, creative potential has proven to be considerably more complex than was expected. We did not expect to find a single, universal key to successful creative thinking, but we were not prepared to find so many facets to the subject.

Considering first the interpreted psychological factors found in the analysis of tests with meaningful, verbal content, the basic traits most clearly related logically to creative thinking include three fluency factors, two flexibility factors, and an elaboration factor. All of these factors are recognized as falling in the same psychological category of divergent-production abilities. They have in common the fact that in measurement of these abilities the tests present the examinee with a certain item, or items, of information and he is to generate from the given information some other items of information. In every case, multiple responses are called for, all different. In many cases, the produced items were probably never previously learned in association with the given information, for example, in a test calling for the activity of forming titles to apply to a given story plot. Information thus comes from memory storage but out of its stored contexts; in other words, by way of transfer. I have elsewhere elaborated on a transfer theory of creative thinking.[1]

The fluency factors pertain to *efficiency* of recall under these general circumstances. They differ from one another in terms of the kind of psychological product involved. Ideational fluency pertains to the rapid generation of units of verbal or semantic information – single ideas or units of thought, such as words to fit a described class of meanings. Associational fluency pertains to the rapid generation of semantic units to fulfill a relationship, having given a specified relation (e.g., a relation of similarity, opposition, part-whole, etc.) and a given unit. Expressional fluency pertains to the production of connected discourse, in the form of phrases and sentences. The psychological product involved is recognized as a semantic system.

Flexibility factors have to do with the *lability* or *fluidity* of stored information. The factor first recognized as spontaneous flexibility is now thought to be a matter of flexibility with regard to classes of information. In tests of this kind of flexibility, the examinee must shift his responses readily from one class to another. It has been said that some uncreative people are suffering from a disease known as hardening of the categories. Such a malady may be a matter of low degree of spontaneous flexibility.

The factor first called "originality" was later recognized as a kind of flexibility, described as *adaptive* flexibility. In one test, the examinee who suggests a relatively large number of clever plot titles earns a high score for this factor. To the fable of the fox and the grapes, for example, if he gives the titles "The fox griped about grapes," or "The sweet grapes turned sour," his titles are regarded as clever because he has produced what we call transformations. A transformation is a change or alteration in meaning or interpretation; a redefinition of some kind. Turning old interpretations of information into new ones makes possible new and different uses of what one knows.

Elaboration means building upon given information to round out a structure, to make it more detailed, or to extrapolate in new directions. Proceeding naturally from what is already given is a matter of producing implications. The given information suggests the first step to be taken, and each completed step helps to determine the following ones. The ability to produce a variety of implications is also in the category of divergent production.

We have just seen that the six divergent-production abilities dealing with meaningful, verbal information differ only in the kinds of products of information involved -- units, classes, relations, systems, transformations, and implications. These are the kinds of psychological products mentioned early in this paper. They are not restricted to divergent production, but have very general application to intellectual abilities in all other operation categories -- cognition, memory, convergent production, and evaluation. It is the relative variety and novelty of the products found in divergent production that links this category of abilities logically with creativity. When we have learned how new products come into existence, how familiar products become transformed, and how known products enter into new connections, we shall know a great deal about creative thinking.

I have mentioned only six divergent-production factors, all pertaining to verbal information. Most likely there are other sets of six, nonverbal, divergent-production abilities parallel to them. One set pertains to (visual) figural information, of which five have been demonstrated. A second set pertains to symbolic information, of which five have also been demonstrated. Thus we have the important general inference that successful creative performance, to the extent that it depends upon aptitudes, is not equally promising in all fields of information. Potential for creative production in the arts is not the same thing as that in mathematics or that in writing, to say nothing of potential for creative handling of problems involving human relations.

It would be incorrect to say that only the divergent-production abilities contribute to success in creative performances. It appears that in the other four operation categories in the structure of intellect[2] we shall find sets of abilities parallel to those in the operation category of divergent production. Thus, in each operation area there is a set of transformation abilities which may contribute to creative potential, particularly in the areas of cognition and convergent production. Convergent production is a matter of generating a single right answer to given information, but it is sometimes necessary to do some searching around or to achieve a transformation in order to arrive at that answer.

In our initial study of creative abilities, we hypothesized and we found a factor interpreted as sensitivity to problems, a factor that has been repeatedly verified. In terms of placement in the structure of intellect, this factor has been recognized as an ability to evaluate implications; for example, seeing the faults in the working of a common appliance or of a social institution. The discovery of a category of evaluative abilities has opened many new possibilities of thinking about thinking. It should have been obvious before that we evaluate just about everything we perceive, or know, or do. Observers of the creative thinking of experts commonly report a terminal period of verification, which, in part, means evaluation. Actually, evaluative operations may occur, and probably do, at any step of the way in thinking, wherever we can gain any feedback information.

Reflection upon the total event of a creative production of some degree of complexity shows how similar that total process is to the commonly held picture of complex problem-solving. The category of problem-solving even applies to the production of objects of art, for the psychoanalysts tell us that in working toward his product, the artist is attempting to solve a personal problem by externalizing a phantasy. The amount of creativity involved in solving any kind of problem is proportional to the amount of novel production the problem solver shows or to the degree of novelty in his solution.

Creative problem-solving involves considerable trial-and-error behavior. Except for instances in which the given information in the problem is sufficient to lead directly to one right answer, every trial is an act of divergent production and the seeing of an error is an act of evaluation. Thus, trial-and-error behavior is an interplay of divergent production and evaluation, ending when some product is evaluated as acceptable. The thinker who can vary his potential solutions easily and extensively has an advantage in getting around to suitable solutions. In everyday life, most problems have no one right answer; there are many answers, a number of which are more or less suitable.

The more able the individual, the more he can produce trial information of high quality, with a higher probability of success. His over-all indulgence in trial-and-error thinking is therefore relatively less. But in view of the necessity for at least some trial and error, and the amount of divergent production that this requires, we should give much more attention to facility in this kind of mental activity and its assessment. Certainly, the person who can think of only one possible answer to a problem is much less likely to solve it than the person who can think of several. A number of investigators have found that tests of divergent production account for many cases of so-called over-achieving, when achievement is assessed in terms of standard achievement tests. The same kind of tests may also be found to account for many cases of under-achievement.

Standard tests and scales of intelligence have practically nothing to contribute to assessment of divergent-production abilities. For example, factor analyses of the Wechsler scales have thus far failed to show any such relationships. Recently I indulged in some armchair analysis of the 140 tests in the latest Stanford revision of the Binet scale, from which comes the impression that, in spite of the great surface variety of the tests, there is an overwhelming weighting with tests of cognitive abilities, with some attention to a few of the factors of memory, convergent production, and evaluation. Only five of the 140 tests appear to offer any appreciable divergent-production variance.

Current, standard academic-aptitude tests have been constituted along the lines of intelligence scales. They therefore emphasize much the same factors and neglect the great majority of the intellectual abilities. It cannot very well be claimed that the neglected abilities are of little importance, for there has been little investigation of the relation of most of those abilities to performance in or out of educational institutions. Many of the neglected abilities should be potentially important in connection with various academic subjects and various occupational activities. For example, some recent data that I have [illegible] show that scores from a few divergent-production tests are correlated with criteria of creative performance of public relations personnel and of advertising copy writers,[8] with coefficients in the range from .60 to .85. The criteria of creativity were essentially rankings of personnel by their superiors.

What is the implication of all this for assessment of academic aptitude? Not being as close to this problem as many of you are, I am not as well prepared to draw up an extensive list of implications. I will mention two or three things that I think should be done.

Believing in proceeding on a foundation of research findings, I would urge that extensive efforts be made to discover the phases of academic learning to which each of the basic intellectual aptitudes is relevant and which of the aptitudes are important to development in each academic subject. At USC we are engaged in a two-year investigation of potentially relevant factors for learning ninth-grade algebra. We are planning to follow this with similar studies of other school subjects.

Although some of the neglected intellectual abilities are quite possibly touched upon in current achievement tests, this is on an incidental basis. As rapidly as research findings become available on the relevance of new aptitude factors, there should be appropriate revisions of academic-aptitude batteries. Present batteries should be examined for redundancies, of which I am sure there are a number, replacing parts with measures of new relevant factors. It is well known that current batteries are by no means equally predictive of achievement in all curricula and in all courses of study. Achievement tests may touch upon many of the deficiencies of aptitude batteries, but they cannot do justice to them.

In introducing tests of divergent-production abilities, some technical difficulties will be encountered. Thus far, we have failed to find machine-scorable aptitude tests of these abilities. It is doubtful that any such tests exist, but with application of ingenuity, something may be done about it. Outside the divergent-production area we have employed a few completion tests also, but it is probable that they can be replaced with machine-scorable tests.

Some of you may be alarmed at the possible large number of intellectual factors that may need to be given attention in the assessment of academic aptitude. I suggest that we save such misgivings until we know how many factors are relevant and how much weight needs to be given to them.

Until such time as measures of creative aspects of aptitude can be included in aptitude batteries, something can be done in the way of gathering information regarding creative promise and in the way of predicting creative production of students. Such information could be used less formally than aptitude scores are used and includes certain biographical signs of creative promise, such as self-initiated or spontaneous extracurricular activities in the form of scientific investigation, writing, or composing. The use of nonaptitude measures is another possibility. MacKinnon and his group at the University of California have shown that the most creative architects and scientists are distinguished by means of interest and temperament inventories.[4] The information they have concerns concurrent validity; presumably predictive validity is also promising.

Thus, although immediate prospects of selecting and placing students with superior creative potential are not certain, future prospects, with the possibilities of using various kinds of information, including aptitude measures, appear to be reasonably bright. The least that we should aim to do is to determine what oversights exist in present aptitude batteries.

REFERENCES

1. Guilford, J. P. "Some New Views of Creativity." In H. Helson, (Ed.), *Theories and Data in Psychology*. Princeton, N. J.: D. Van Nostrand Co., 1963 (in press).
2. Guilford, J. P. and Merrifield, P. R. "The Structure-of-Intellect Model: Its Uses and Implications," *Reports from the Psychological Laboratory*, No. 24. Los Angeles: University of Southern California, 1960.
3. Elliott, J. M. "Advancing our Methods for Measuring Creative Abilities." *Information Service for A.N.A. Members*. New York: Association of National Advertisers, May 1962.
4. MacKinnon, D. W. "Fostering Creativity in Students of Engineering." *Journal of Engineering Education*, 1961, 52, 129-142.

Some Methodological Issues in the Study of Creativity

Robert L. Thorndike,
*Teachers College,
Columbia University*

ONE FEELS put in a somewhat uncomfortable, perhaps even untenable position when called upon to look critically at the research on creativity. As one reads the reports in this field one encounters frequent references to the importance for creativity of a complete openness to any idea, and of the lethal impact upon the creative act of an evaluative and critical approach to life. So from this point of view, the critic is condemned before he starts to being at best noncreative and at worst anticreative — in the physicist's sense of antimatter — that is, to being a force that destroys creativity. But if this be true, I must just make the best of it, and trust that this will be a creative critique, so that it won't destroy creativity, but only call attention to some possible directions for creativity research.

For the moment, let us forego any formal definition of creativity. We may need to back into the issue of definition as we proceed further with the problem of evaluation.

As a point of departure, let us raise the question of whether research should focus on the creative act or the creative individual. From some points of view, the specific act or unitary sequence of acts is the attractive unit of study. Thus, we might get protocols from the poet as he writes a poem, or from the engineer as he plans a device permanently to eliminate leaky faucets. We might examine the paintings produced by a class of first-graders, or the compositions written by high school seniors. But how are we to determine when a creative act has taken place? Is creativity a matter of kind or degree? Do we identify creativity through the process or the product?

The product is, of course, much more convenient to study. It is there and can be examined at leisure. We can replicate judgments so as to get the needed degree of reliability—though the replication may not do much to improve the validity of the resulting judgments. We can divorce our evaluation of the product from possibly irrelevant attributes of the producer, such as his or her personal charm, his or her group membership.

But can we judge creativity from a product? We may be able to appraise excellence in a product, by some implicit or explicit criteria of excellence. We can almost certainly appraise unusualness of a product, if that appeals to us as an indicator of creativity. And we can certainly very readily appraise sheer voluminousness of production, which sometimes seems to be offered in lieu of excellence. However, the production of even a good product need not constitute a highly creative act for the individual. What proportion of the stories of even the most gifted raconteur are new creations as he tells them, or are even of his own original devising? In every field, the skilled and well-trained workman may turn out products that are technically excellent, but that involve little of what we would consider to be creativity in the producing.

And even the statistically rare need not be creative as far as the specific individual is concerned. Unless we know his individual history, we cannot be sure but that he is feeding back to us rather mechanically something he saw or heard only a day or two before. Before we can make any sure appraisal of the creativity of a production by an individual, we need to know intimately the microstructure of past experience of that individual so that we may evaluate the product against the background of that past experience. Whether an act is creative, in the sense of representing a novel synthesis *for that individual*, depends not only upon the product produced, but also on the past history of the producer. This holds true equally of all tasks that aspire to appraise "thinking" or "problem-solving" on the part of an individual. One man's problem solution is another man's habituated response.

It would seem then that we can appraise the unusualness of a product, or the excellence of the product in terms of certain criteria; but it is questionable whether we can appraise the creativity of a specific product. Creativity lies in the relationship between the product and the producer. To study creativity in an act, it seems that we must return to the process and study that process.

There have been a number of attempts to study the creative process over the years. Some have been based on retrospective reports by particularly successful creators or problem-solvers, some upon concurrent observation of somewhat lesser souls. These have brought out a number of recurrent features in creation and problem solution – the need for intense immersion in a problem, the role of preconscious activity during a period of withdrawal from the task, the frequently sudden redirection and restructuring of perceptions, and the like. Do the features frequently described in the creative process give us a basis for categorizing acts as creative or noncreative, or for scaling them with respect to degree of creativity?

If we were to combine features of observable behavior with concurrent verbalization of his mental content by the subject, and add to this a retrospective self-analysis by the subject, we could probably separate those acts that called for only a routine application of available knowledge or skill from those that called forth a genuinely new synthesis of experience. But we could not do this perfectly nor completely. It might even be possible to achieve some consistency in sorting acts into several categories of an ordered series with respect to a dimension of automatic and routine vs. productive and novel. For the person who is interested in studying the situational factors influencing creativity or the effect of different kinds of experiences or training upon the creativity of behavior, such a crude scaling might have value. Thus, it might be possible to distinguish those effects of experience which lead to effective routinizing of certain skills, such as applying the congruency theorems of geometry or using Roget's *Thesaurus* in writing limericks, from those which lead to a freer and more spontaneous attack upon novel problems. It might be possible to judge whether the character of the individual's mental processes changed as a result of the training, or whether the training merely provided a larger foundational background of knowledge or skill to which essentially the same processes were applied. It might at least make it possible to weed out from among the products of an individual – which we might then want to evaluate for their excellence – those that were routine reproductions of habituated responses. However, an enterprise of this sort would be extraordinarily laborious. Whether the yield would justify the labor may well be questioned.

Let us turn our attention now from the act to the person. The question now centers on whether it is meaningful to attempt to identify individual differences in creativity among persons, and whether "creative," used as a rather global term to characterize a person, has meaning.

Once again, it would be possible, in theory at least, to describe differences between persons either in terms of process or of product. Is the creative person to be identified as the person who frequently engages in creative acts – acts in which he restructures and recombines the components of his experience, producing syntheses which are novel to him and compatible with the demands of the situation? Or is the creative person to be identified as the individual who frequently produces good products – good art, science, or social invention – good in the sense of being approved or valued by society?

In the case of the person, I suspect that we will rarely find it feasible to evaluate creativity in terms of process. It hardly seems practical to get the type of evidence about individuals that would give us an adequate basis for judging individual differences in frequency or quality of creative acts. Our appraisal is almost certain to fall back in large part on the products the individual has produced, and be a judgment of some quality in these products. The products could, of course, be those produced within a circumscribed time and set of circumstances – that is, under essentially test conditions, as when an essay is written on an examination or a defined test of "fluency" or "originality" is administered. But here, as elsewhere, we tend to think of tests as predictors of, rather than criteria of life performance. What investigators have often relied upon is some type of global evaluation of real-life productions. Persons knowledgeable in the field, be it one of the arts, one of the sciences, or one of the professions that are neither art nor science, are asked to nominate or rate colleagues with respect to their contribution to the field.

There is no question that with some degree of agreement specialists can appraise individuals in their own field with respect to some composite of excellence. Individuals differ in the reputation that they have with their colleagues. Some are more favorably perceived than others. Certainly, if psychologists, or any other group, were asked to rate one another on creativity, there would be some nucleus of agreement. The stickier question is whether colleagues could discriminate "creativity" from "productivity," "visibility," or "general reputation." Do ratings and nominations for "creativity" only have a certain amount of general validity as indicators of professional reputation, or do they have *specific, differential* validity in picking out those who are creative from among all those who are professionally successful and visible? What would be the correlation, corrected for rater unreliability, between ratings for creativity and ratings for some other indicator of general reputation? Our general experience of the halo in ratings does not make us too optimistic about the prospects for a valid differential rating.

When we use any term to designate an attribute of an individual, be it abstract intelligence, sociability, or creativity, the distinct label implies that there exists a set of behaviors of the individual that can be grouped together because they exhibit a common quality. Furthermore, if the label is a useful one, the set of behaviors that it designates can be differentiated from the sets designated by other labels. Thus, we expect the sociable person to choose to be with other people and to express pleasure in being with them. We distinguish this set of behaviors from the set to which we apply the term "ascendant" because the sociable person will in many cases show no tendency to take a leadership role, but will be content with merely being in and a member of a group.

We may appropriately ask how well the attribute "creativity" meets these joint criteria of designating a reasonably extensive set of behaviors which (1) have some degree of coherence and (2) can be distinguished from other sets of behaviors.

The question may be asked in relation to the activities of life. The question will then be phrased: Does the person who behaves creatively in one type of life situation tend also to show creative behavior in others? To what extent is the inventive scientist also a creative amateur in music, art, or writing? To what extent does the original architect also exhibit ingenuity and improvisation in interpersonal relations?

To obtain dependable evidence on such life performances would certainly be extraordinarily difficult. As a matter of fact, in many fields it is certainly difficult to make diagnostic judgments that discriminate between a *creative* performance and a *competent* one. And it is perhaps unreasonable to expect comparable performance in different fields. It could be argued with some plausibility that a person who finds an outlet for his creativity in one direction may have no need to show creativity in others. If this be so, then either we must apply some modifier to the label "creative" – i.e., creative in art, creative in physical science, creative in corporate finance, creative in home decorating – thus specifying more limited sub-sets, or we must unite these sub-sets on the basis of the correlates of their creative behavior.

A good deal of the research on individuals identified by some type of nominating procedure as being creative has focussed upon the characteristics of these creative individuals, or in other words, upon the correlates of creativity. Thus, this has been one main focus of effort for MacKinnon and his associates at the California Institute of Personality Assessment. They report, for example, that on the Barron-Welsh Art Judgment Test not only outstanding architects but also outstanding writers, scientists, and mathematicians tend to choose complex, asymmetrical and "messy" line drawings in preference to simple, symmetrical, and (judging from the illustrative examples) "sterile" ones. A number of other uniformities are reported for groups selected as leaders in different fields, and become the basis for affirming a degree of generality to creativity. These uniformities arise from the assessment approach that is espoused by the California group. The comprehensive assessment is clearly fruitful in suggesting hypotheses as to correlates of creativity, but here perhaps as dramatically as anywhere the limitations of the assessment approach as a design for *testing* hypotheses become apparent. As Gough pointed out, in developing some multiple correlations in the high 90's with rating criteria on a group of engineers, with over 100 items of data for a sample of 45 cases some mighty dramatic relationships can be generated. The critical issue is whether the relationships will be maintained in a new sample of cases. And an adequate stock of new cross-validation cases is difficult to come by when the persons are such rare birds as some of the California groups of highly evaluated architects, writers, and mathematicians. We face something of a dilemma. We are loath to accept lesser degrees of talent in order to get numerically convincing samples, and yet we feel the pressure for verification of reported findings. The value conflict between the hypothesis generator and the hypothesis tester is in some ways parallel to the discrepancy between the perceptual and judgmental modes which are alleged to characterize the more and the less creative. Hypothesis testing certainly has less of a flavor of creativity, but still has an essential role in the fabric of science.

In the test domain, as distinct from the life-activities domain, the question as to the meaningfulness of a general rubric of "creativity" can be raised somewhat more incisively. We may ask whether there is a variety of different test behaviors that (1) seem reasonably to pertain to the concept of "creativity," (2) are associated so that a person who tends to exhibit one also tends to exhibit the other, and (3) are distinct from other sets of test behaviors such as the set to which we have applied the term "abstract intelligence."

The issue we raise is closely analogous to the issue that has been raised with respect to the concept of intelligence, around which so much research and controversy have centered. Though ways of interpreting the facts in the domain of intelligence have differed, there would be general agreement as to the basic facts, I believe. The essential points are that although there is a degree of specialization of intellectual functioning, so that tests within a specific region of content or process correlate more highly than those from different regions, still the correlations across regions are appreciably positive. It is these uniformly positive correlations, whether conceptualized as G, or a second-order factor among the primary factors, or as an overlapping of group factors, that give some substance to the general concept of abstract intelligence and some reasonableness to pooling a set of sub-tests into a common score.

We may appropriately ask whether there is another broad second-order factor in the test domain, distinct from the traditional G, to which the term "creativity" can appropriately be applied. The existence of such a distinct factor is strongly implied in the publications by Getzels and Jackson and by Torrance, among others, and Guilford and his associates have fairly sharply differentiated between tests of convergent and divergent thinking. How well does this differentiation of two broad cognitive domains hold up in practice?

I have re-analyzed some of the published data to try to get a partial answer to these questions. The first set of data I examined were the correlations reported in Table 1 of Getzels and Jackson's recent book *Creativity and Intelligence*. Getzels and Jackson emphasize the lack of correlation between the traditional intelligence test and the measures that they used to appraise creativity. However, the intercorrelations of the five "creativity" tests were themselves not very high. It is of some interest to extract a first factor from this table of correlations and compare the factor loadings of the several tests.

Table 1

First Factor Loadings of Getzels and Jackson's "Creativity" Tests and IQ

	Boys	Girls
1. Word Association	.69	.70
2. Uses	.47	.48
3. Hidden Shapes	.58	.60
4. Fables	.41	.42
5. Make-up Problems	.58	.72
6. I. Q.	.52	.50
Average	.54	.57

The results are shown in Table 1. Thus, we see that on the first factor common to these six measures, the factor loadings are all fairly modest and the loading for the conventional intelligence test falls about midway among the "creativity" tests.

However, these data are less than ideal for answering the question of the factor structure of the intelligence and creativity domains because they include only one undifferentiated measure of intelligence. More comprehensive sets of data are provided in some of the studies by Guilford and his associates. I have re-analyzed certain of these by somewhat primitive methods to explore the point at issue. In general, what I have done is to take the two tests with highest factor loadings on a given factor to represent that factor. Partly on personal judgment and partly from the discussion provided by the authors, I have classified the factors as old-line or convergent thinking factors or as new-type divergent thinking (i.e., "creativity") factors. I have prepared a miniature correlation table, made up of just these pairs of tests, for each factor, re-arranged so as to appear in the two designated broad clusters. Then I have found the average correlation for each test, first with the tests of other factors within its own domain and then with all the tests in the other domain. The first correlation provides an indication of the extent to which a test "hangs together" with its own domain, and the second an indication of the extent to which it overlaps the other domain, or the extent to which the domains themselves are indistinguishable.

In Table 2, you see the results for this analysis applied to a factor analysis of the domain of verbal fluency carried out by Guilford and Christensen. This study appeared to yield three factors that would be considered old-line convergent factors — to wit, Verbal Comprehension, General Reasoning, and Eduction-Verbal Correlation. Five new-style divergent factors appeared to emerge, labelled Originality, Expressional Fluency, Ideational Fluency, Word Fluency, and Associational Fluency. The average correlations are shown in the body of the Table. Thus, the across-factors correlation for the old-line tests was .435, while for the new-type tests it was .267. The average across-domains correlation was .238. The Word Fluency and

Table 2

Average Correlation Across Factors Fluency Study

Guilford and Christensen

Divergent Factors

Factor	Test No.	Average Correlation			
		Divergent		Convergent	
Originality	35	.155	.255	.196	.245
	36	.355		.294	
Expressional Fluency	24	.303	.270	.174	.198
	28	.238		.221	
Ideational Fluency	17	.340	.265	.260	.158
	19	.190		.057	
Word Fluency	2	.228	.250	.260	.290
	5	.271		.320	
Associational Fluency	7	.239	.294	.286	.300
	22	.350		.314	
Average		.267		.238	

Convergent Factors

Factor	Test No.	Average Correlation			
		Divergent		Convergent	
Verbal Comprehension	15	.263	.236	.520	.478
	38	.209		.436	
General Reasoning	34	.274	.202	.510	.398
	40	.129		.286	
Eduction	11	.289	.276	.466	.439
Verbal Correlation	12	.263		.412	
Average		.238		.435	

	Convergent Tests	Divergent Tests
Factor G	.66	.36
Factor D-C	–	.37

Associational Fluency tests correlated slightly higher with the convergent than with the rest of the divergent tests.

If we conceptualize these findings in terms of an orthogonal two-factor structure, assuming that what is common to the old-line tests is some sort of a G factor (here pretty heavily saturated with a verbal component), that what is common to the old-line and new-style tests is entirely this same G, and that whatever more there is tying together the new-style tests is our D-C (divergent or creativity) factor, we get the factor structure shown at the bottom of Table 2 as representing the average test in each of the two clusters. That is, the average old-line test has a loading of .66 or G, which accounts for 43½ per cent of the test variance. The average new-style test has loadings of .36 on G and .37 on D-C, accounting for roughly 13 and 14 per cents of variance, respectively. Our second broad factor has reality, but the tests in its domain are only about one-third as saturated with it as the old-line tests are saturated with G.

A second analysis was made based on the factor analysis of creative-thinking abilities by Wilson, Guilford, and others. This was a rather more extensive battery, providing six presumably old-line and eight new-style factors. The results of the initial analysis are shown in Table 3. Both broad factors appeared more weakly in this set of data, perhaps because the content of the test materials ranged more widely. However, the old-line cluster shows an average cross-factors correlation of .23, and the new-style cluster an average of .14. The average cross-clusters correlation was .12. Again, conceptualizing this in a two-factor structure, we find a G factor of .48 for the old-line tests, while the new-style tests have average loadings of .25 on G and .28 on D-C. Though all loadings are reduced, the relative values are quite comparable to those for the other study. What is common among the divergent thinking or creativity tests can be divided about equally between the conventional cognitive factor and some other factor that does not appear in the old-line tests. This new factor accounts for about a third as much of the common variance of these tests as the G factor does for the conventional tests.

A further examination of Table 3 suggested that certain of

Table 3

Average Correlation Across Factors

Wilson, Guilford, and others

Divergent Factors

Factor	Test No.	Average Correlation: Divergent		Average Correlation: Convergent	
Word Fluency	53	.135	.126	.268	.250
	28	.118		.232	
Associational Fluency	39	.204	.190	.125	.158
	11	.175		.190	
Ideational Fluency	6	.098	.143	—.025	.003
	8	.188		.031	
Originality	7	.159	.103	.059	.006
	19	.047		—.046	
Adaptive Flexibility	16	.042	.070	.250	.234
	17	.097		.218	
Spontaneous Flexibility	10	.155	.188	.089	.086
	22	.220		.083	
Redefinition	29	.062	.102	.184	.139
	31	.141		.094	
Sensitivity to Problems	24	.173	.154	.075	.052
	25	.134		.030	
Average		.14		.12	

Convergent Factors

Factor	Test No.	Average Correlation: Divergent		Average Correlation: Convergent	
Verbal	33	.154	.128	.274	.231
	43	.102		.188	
Numerical	50	.136	.135	.242	.269
	52	.134		.296	
Perceptual	38	.102	.109	.188	.223
	47	.116		.258	
Visualizing	35	.101	.076	.268	.229
	51	.052		.190	
Reasoning	34	.146	.136	.284	.260
	44	.125		.236	
Closure	37	.067		.140	
Average		.12		.23	

	Convergent Tests	Divergent Tests
Factor G	.48	.25
Factor D-C	—	.28

Robert L. Thorndike

the factors might be misplaced. The Word Fluency and Adaptive Flexibility Tests showed markedly stronger affiliations with the old-line than with the new-style tests. Perhaps the limits of the clusters should be redefined. This has been done in Table 4, by shifting the Word Fluency and Adaptive Flexibility tests to the old-line convergent cluster. The effect has been to increase appreciably both the homogeneity and independence of the remaining cluster. The average within-cluster correlation rises to .185, while the between-clusters correlation drops to .077. The factor loadings for these redefined clusters are shown at the bottom of the Table. The reduction in the G loading and increase in the D-C loading for the smaller group of new-style tests can be noted. The G factor now accounts for only a little over two per cent of the variance, while the D-C factor accounts for 16 per cent – about two-thirds as much variance as the G factor accounts for in the old-line tests.

I would very much like to apply this same type of critical (if not creative) analysis to the tests that Torrance has been developing at Minnesota. So far, I have not encountered a set of data that lent themselves to this approach. Though Torrance has expressed commendable concern about rater reliability in appraising the protocols from his tests, I have not encountered the same type of concern about trait reliability, that is, the consistency with which his tests measure some common attribute to which a common designation may legitimately be applied. Though Torrance specifically disavows intending to produce a test to produce a Creativity Quotient that would constitute a characterization of an individual, he often uses a team of his tests as if they did produce one, or at least as if they had enough in common to justify pooling them into a single composite score. I would suggest that a good deal of further study of the behavior domain is needed before this is done.

The analyses that I have reported to you here suggest that there is some reality to a broad domain, distinct from the domain of the conventional intelligence test, to which the designation "divergent thinking" or "creative thinking" might legitimately be applied. However, they suggest that this is

Table 4

Modified Average Correlation Across Factors

Wilson, Guilford, and others

Divergent Factors

Factor	Test No.	Average Correlation			
		Divergent		Convergent	
Associational	11	.163	.180	.194	.172
Fluency	39	.198		.151	
Ideational	6	.136	.188	−.020	.007
Fluency	8	.241		.034	
Originality	7	.178	.123	.071	.017
	19	.068		−.037	
Spontaneous	10	.177	.230	.090	.082
Flexibility	22	.282		.074	
Redefinition	29	.030	.107	.175	.125
	31	.184		.075	
Sensitivity	24	.202	.182	.080	.059
to Problems	25	.161		.038	
Average		.185		.077	

Convergent Factors

Factor	Test No.	Average Correlation			
		Divergent		Convergent	
Verbal	33	.087	.099	.299	.242
	43	.111		.186	
Numerical	50	.082	.076	.260	.283
	52	.070		.306	
Perceptual	38	.067	.074	.194	.220
	47	.081		.246	
Visualizing	35	.054	.036	.260	219
	51	.019		.178	
Reasoning	34	.105	.102	.280	.252
	44	.100		.225	
Closure	37	.044		.149	
Word Fluency	28	.108	.121	.232	.246
	53	.134		.259	
Adaptive	16	.020	.049	.240	.232
Flexibility	17	.078		.224	
Average		.077		.235	

	Convergent Tests	Divergent Tests
Factor G	.485	.156
Factor D-C	–	.401

a rather more nebulous and loosely formed domain than that of conventional intellect. If this be true, an adequate domain measure, which clearly brings out this general characteristic relatively unencumbered by variance due to error or to specific factors, will be hard to come by. Different creativity measures will be less equivalent and interchangeable than different intelligence measures. We will be well advised to use the term "test of creativity" with even more circumspection than we are learning to use in speaking of "test of intelligence." But at the same time that we urge a considerable degree of caution, we can all applaud the attempts to develop a new and potentially fruitful field of appraisal.

Session II

Theme: Comparability of Scores

Can Useful General-Purpose Equivalency Tables Be Prepared for Different College Admissions Tests?

WILLIAM H. ANGOFF,
Educational Testing Service

IN RECENT MONTHS complaints have been heard from various sources that the amount of testing that is going on these days is excessive and burdensome to the schools. These complaints are primarily directed at the kind of testing that has been called "external," by which is meant the testing that is demanded of students and schools by agencies external to the schools – e.g., colleges requiring standardized information regarding the abilities of students who are applying either for college entrance or for college scholarships. One solution that has been offered for this excessive testing calls for the computation of equivalency tables which would allow the conversion of scores obtained on the test of one publisher to the score scale of another publisher. A grander solution that has been suggested is to convert the scores on tests produced by all publishers to a single, common score scale. These conversions would presumably make it possible for a school to administer to its students only one test and then to submit to each college admissions officer or scholarship agency a converted score reported in terms of the score scale that is considered acceptable by that agency. In this way, it is claimed, the student would be spared the repeated testing that is now ordinarily required of him.

The validity of the claim that external testing today is excessive has been examined thoroughly, I am sure, and I doubt that I could add anything significant here to that discussion. However, you may be interested to know, if you do not already know, of an excellent article written by Frank B. Womer of the University of Michigan for the Fall 1961, issue of the *North Central Association Quarterly*, in which he discusses the "Pros and Cons of External Testing" thoroughly and quite objectively, I believe. You may also know of a study conducted by the College Board in 1960 in which the use of tests in the eleventh and twelfth grades was surveyed. However, aside from the question that external testing is excessive, what I would like to consider here is the adequacy of the proposed solution — of equivalency tables that would relate the tests of all publishers.

This solution has some obvious attractions. It also has some objectionable features which may be summarized by saying that such tests will not necessarily be parallel forms — that is, they will not necessarily be measuring precisely the same function — and that in such instances it is not possible to provide a unique conversion that will be appropriate for all groups who are likely to take the tests or for all purposes for which the tests are likely to be given. It is of course reasonable to equate the tests of different publishers when these tests are known to be parallel. And quite possibly some *are*. On the other hand, it is more than likely that in the majority of instances they are not, since they were not constructed with that intention in the first place.

Before going on to consider the problems of equating tests of different function, it may be helpful to say what we mean when we talk about the equating of scores on tests of the *same* function. In this case, in the conversion of scores from one form to a parallel form of a test, say from one form of the ACE Psychological Examination to another, or from one form of the College Board SAT to another, we have simply the problem of transforming the system of units. We ordinarily consider this problem as one that is directly analogous to the problem of conversion from centimeters to inches, from Centigrade to Fahrenheit, from pounds to grams, and so on. On the other hand, the question of converting scores across tests of *different* function is another matter. Here we are concerned with the problems of conversion from one kind of dimension to something quite different — from the scale of inches to the scale of pounds, for example. This is not necessarily an extreme or unreasonable example. For instance, the correlation between height and weight for adults of the same age is in the region of the low .50's. The correlation between observed scores on two tests of mathematical ability — SAT-Mathematical and ACE-Quantitative, for example — both of which, it should be pointed out, were designed to be predictive of scholastic success in the quantitative areas, is in the low .60's. The true-score correlation between these tests is not much higher — around .70. On the face of it, it seems unreasonable to ask for a conversion from inches to pounds. Why then, is it not similarly unreasonable (or at least *almost* as unreasonable) to ask for a conversion from SAT-M to ACE-Q? Yet requests for this conversion are not at all unusual.

Even in the conversion of scores across tests of the *same* function we can enumerate at least four sources of error. One source of error lies in the unreliability of the measuring instruments themselves, and affects the stability of statistics obtained in the collection of data. A second source of error lies in the design of the equating experiment and the method of treating the data: The various methods of equating are not equally reliable. A third source of error lies in the choice of samples used to establish the conversion line. Different samples will yield lines that differ randomly from one another. Finally, there is a source of error that is characteristic of any statistical enterprise — the error associated with the size of the study sample. Obviously, larger samples will yield more reliable data.

Nevertheless, in spite of all these kinds of error, it is reasonable, in the equating of parallel forms of a test, to postulate a unique "true" line and to consider that errors resulting from the unreliability of tests, from design and method, and from the sampling of people only cause random departures from this true line. On the other hand, the problem of "comparable" scores, that is, the conversion of scores of *different* function, has additional complexities. Here, as in the physical analogy in which we convert height to weight, the request for equivalent scores rests on an inappropriate premise — that there is *a single* conversion system — a unique system that exists and needs only to be estimated with the use of appropriate data. While it is certainly true that one can estimate or predict by means of a regression equation the most likely weight of an individual of a given height, or predict scores on the SAT from knowledge of scores on the ACE, in neither case can the prediction be considered simply a transformation of unit systems. Furthermore, just as regression systems will differ, depending on various considerations like the conditions of measurement and the characteristics of the group measured, so will conversion lines which purport to translate scores across tests of different function differ – and they will differ systematically, that is, in a predictable fashion, depending on certain known considerations.

For one thing, different conversion lines will result, depending on the particular definition of comparability employed. For example, one might ask what the most likely score on one test would be, given a particular score on the other. This would be one definition. Or one might equate standard-score deviates on the two tests. Or one might equate equivalent percentiles. These would be other definitions. Or one might, for example, equate scores that yielded equal predicted scores on one criterion or another. Here the use of different prediction criteria might yield different conversions. There is probably an unlimited number of definitions, each one conceivably yielding a different conversion. In contrast to this multiplicity of conversions for nonparallel tests, all methods or definitions of equivalent scores for *parallel* tests should lead to the same conversion equation – that is, except for random error of the kind described before.

Secondly, different lines will result when different populations are used to form the basis for deriving the tables of comparable scores. For example, there will be one conversion line between Verbal and Mathematical test scores at the college entrance level for boys and quite another for girls. This fact is immediately apparent if one examines the mean scores of boys and girls on the Verbal and Mathematical sections of the SAT. Girls characteristically score slightly higher than boys on Verbal but considerably lower than boys on Mathematical. The conversion line between Verbal and Mathematical scores when derived from the data for boys would be much different from the conversion line derived from the data for girls. As another illustration of the same general point we might consider some data collected by Dr. Marjorie Olsen in a study in which an attempt was made to predict Graduate Record Aptitude Test scores from College Board SAT scores earned three and four years earlier. In this study it was found that students who later majored in liberal arts scored 40 points higher on SAT-Verbal than students who later majored in engineering. On SAT-Mathematical, however, the picture was considerably different. Here, the students who later majored in liberal arts scored 45 points lower than students who later majored in engineering. The data available on the GRE Aptitude Test four years later showed the same picture, even more pronounced. Here, the students who had majored in liberal arts scored 56 points higher than the engineers on GRE-Verbal. On GRE-Quantitative, however, the engineers scored 76 points higher than the liberal arts students. While these results may not be too surprising, they do point up the fact that any so-called equating between tests of different function, say between SAT-Verbal and SAT-Mathematical or between GRE-Verbal and GRE-Quantitative, would not necessarily yield the same results for different kinds of groups. When we consider the problem from a slightly different point of view, we recognize that conversion lines for tests of different function *must* be different for different groups if the groups show different profiles on the two tests. One might even say that because they reveal group characteristics and are not independent of the groups on which they are based, these kinds of conversions are themselves another way of representing group profiles.

A third reason for the nonuniqueness of comparable scores is the fact that differential selection will have a pronounced and predictable effect on the conversions. Let us suppose, for example, that we collect data on the SAT and the ACE for a group of individuals in order to construct a conversion table from scores on one of these tests to scores on the other, and let us suppose also that these individuals have been explicitly selected on the SAT. The distribution of resulting SAT scores will, of course, be sharply curtailed. The distribution of ACE scores, on the other hand, will be only moderately curtailed. A conversion line derived from these data could be far different from a line derived from a group of unselected students, or from a group of students who were selected on ACE. Even if the students have not been selected on one of the two tests, but on some third variable, say high school grades, their resulting distributions on the two tests and consequently the conversion of scores on the two tests could be considerably altered. This is so because the two score distributions will be unequally curtailed in rough proportion to the correlation of the two tests with the selection variable. The test that correlates more highly with the selection variable is the test on which the distribution of scores will be more drastically affected by the selection. The unequal effect of this selection is what causes the conversion for the selected group to be different from the conversion for the original unselected group. In contrast, scores taken from parallel tests of equal length will be equally affected by selection on a third variable, since they would correlate equally with that variable. The result is that there would be no predictable effect on the conversion equation.

I should emphasize that these problems are inherent in the derivation of a system of comparable scores, that is to say, in the conversion of scores on nonparallel tests, and that they do not replace or preclude the kinds of sources of error that were mentioned before in connection with equivalent scores that are taken from parallel forms. The problems of comparable scores exist *in addition to* the problems of equivalent scores.

By referring to appropriate data, it can be demonstrated quite readily that conversions for nonparallel tests are not unique, but depend in each case on the nature of the group used to establish the conversion. Table 1 shows two sets of conversions between ACE-Quantitative and SAT-Mathematical, one based on a group of 333 male students, the other based on a group of 582 women, both enrolled in the liberal arts curriculum at a large western state university. Clearly, each of the ACE-Q scores listed in the column at the left yields two converted scores that differ from each other appreciably. If it were necessary to convert these two tests to a single "common currency," so to speak, to be used for students of both sexes, there would be considerable doubt as to which one of these two conversions would be the proper one to use. In contrast, Table 2 also shows two sets of conversions for the same two groups of students, this time for tests that correlate more highly, ACE-Linguistic and SAT-Verbal. The results in Table 2 are entirely different from those in Table 1. Here, because the tests are more similar, in terms of what they measure, than are ACE-Quantitative and SAT-Mathematical, the conversions are in much greater agreement.

Table 1

Conversions from ACE-Quantitative to SAT-Mathematical

ACE-Q Scores*	SAT-M Scores		
	Men Liberal Arts	Women Liberal Arts	Difference
6	250	226	24
9	333	299	34
12	416	373	43
15	499	446	53
18	582	519	63
21	665	592	73
24	748	665	83

* Special scale for ACE; Mean = 13, Standard Deviation = 4.

Table 2

Conversions from ACE-Linguistic to SAT-Verbal

ACE-L Scores*	SAT-V Scores		
	Men Liberal Arts	Women Liberal Arts	Difference
6	322	316	6
9	393	383	10
12	464	450	14
15	535	518	17
18	605	585	20
21	676	652	24
24	747	719	28

* Special scale for ACE; Mean = 13, Standard Deviation = 4.

It may give concrete value to the points that have been made here if we constructed a special set of illustrative data. Let us say that we have two college admissions tests (x and y) and that we construct a table for converting scores from the scale of one test to the scale of the other. For the sake of the illustration, let us say that the table was constructed on the basis of data that yielded a mean of 50 and standard deviation of 10 on one test (Test x) and a mean of 500 and standard deviation of 100 on the other test (Test y). The conversion based on these data would comprise the figures in the second column of Table 3. Now let us curtail the lower forty per cent of the distribution of these students on a third test (Test z) — let us say that these forty per cent were rejected for admission at a college — and let us develop a conversion for Tests x and y on the basis of the selected students. Let us say, finally, that Tests z and x correlate .60, and that Tests z and y correlate .95. We can estimate that the sixty per cent of the group who were selected for admission will have a mean of 53.9 and a standard deviation of 8.9 on Test x and a mean of 561 and a standard deviation of 69 on Test y. The conversion based on these data appears in the third column of Table 3. The differences between the two conversions appear in the fourth column. Obviously, these differences are not inconsiderable. Which of these two, then, is the appropriate conversion, and is it appro-

Table 3

Conversions for Imperfectly Correlated Test Scores, Before and After Incidental Selection

X-Scores	Corresponding Y-Scores		
	Before Selection	After Selection	Difference
20	200	298	98
30	300	376	76
40	400	453	53
50	500	531	31
60	600	609	9
70	700	687	13
80	800	764	36

priate under all circumstances? Clearly, neither one is appropriate under all circumstances for the reason that the tests are not parallel forms.

Added to the systematic effects on comparable score conversions that were discussed earlier — namely, the definition of the kind of comparability in question; the choice and size of the group which would be used to establish the conversion; and finally, the related effects of selection — there are still other problems to face when we consider the equating of nonparallel tests. Particularly important are the problems of interpretation. For example, there is the danger that converted scores will be assumed to possess the reliability and validity of the scale of scores to which they are converted. While it should go without saying that scores on one test, however well they may be converted to another scale, retain all the basic characteristics of the original test, many test users will assume that, once converted, they should behave in all respects like the test whose scale they have adopted. Obviously, this assumption is unjustified. It becomes evident that this is not an insignificant concern when we reflect that this kind of misinterpretation favors the publisher of the poorer, less reliable test. And it is probably safe to say that scores on the less reliable test would more often be converted to the scale of the more reliable test than vice versa.

Related to this is the very real worry that the existence of tables of comparable scores will favor the test that is readily available for purchase and is sanctioned by the publisher for administration under less secure and less carefully controlled conditions. Here too it is likely that this test would be the one that is converted to the scale of the test in which only the highest standards of control are permitted in and out of the testing room than vice versa. Since it is also reasonable to assume that the test that is administered under less stringent conditions is the one that is likely to yield the higher score, it is easy to see that the candidate, say, for college admission, who is given the test under these conditions is the one who stands to gain in comparison with a competing candidate who is required to take the more secure test.

Another kind of error occurs in the interpretation of converted scores when regression methods are used for establishing the conversion line between two tests. Predicted scores are regressed, that is, they have a reduced standard deviation; original scores are unregressed. The pooling of these two kinds of data, one taken from actual obtained scores and the other taken from regressed scores, would be improper. Yet it is for the very purpose of pooling such data, or at least, considering them in conjunction, that conversion tables are ordinarily called for. Let us say, for example, that we have a conversion of this sort from ACE to SAT. It is interesting to note here that, to adopt the winning strategy, good students would be well advised to take the test that does not undergo this type of conversion – the SAT, that is – since, by doing so they will avoid the unfavorable effects of regression, in which scores at the upper end of the scale are characteristically depressed. *Their* scores would depend only on their own high level of performance. Poor students, on the other hand, would be well advised to take the ACE, because in the conversion to the SAT their scores would naturally be regressed upward toward the mean. Thus, by following this advice and taking advantage of statistical laws, they manage to achieve a higher score than they deserve – that is, a higher score than would be earned by students of the same ability who are not given this advice. It is important also to note that the lower the correlation between these tests the greater the extent of the regression and the greater the effect of the strategy. The other side of this coin is also worth noting: The good student would be ill advised to take the ACE because his score would also regress in the conversion process, but downward toward the mean.

There are other situations as well where it would be to a student's advantage to take into consideration the effects of regression. Suppose, for example, a table of comparable scores were developed for two tests of dissimilar validity by defining as equivalent those scores that yielded equal predicted scores on some single criterion, say college grades. Since scores on both tests would inevitably regress toward the mean and since the regression would be more pronounced for the less valid test, it is clear that the winning strategy for the poor student would be to take the less valid test, because it is on this test that his scores would undergo the greater upward regression. The opposite strategy would be appropriate for the good student. His scores would necessarily regress downward in the conversion process, but if he were circumspect and took the more valid test he would insure himself against substantial loss due to the effects of regression.

In spite of all that has been said so far, there are times, nevertheless, when it is reasonable and appropriate to derive tables of comparable scores for tests of different function. The methods of deriving these tables are much like, and in most instances identical to, the methods of equating scores on parallel tests. The ideal plan for equating would be one in which we would select a sample of individuals to whom we would administer both Forms x and y and see to it that practice on the first form administered did not affect performance on the second form. However, since the ways of doing this are not immediately apparent, we find it more convenient to divide the group into random halves and give Form x to one half of the group and Form y to the other half of the group. Then, making the assumption that there are no systematic differences between the groups, we equate the means and standard deviations on the two tests to produce a linear equation relating the two sets of scores. This procedure in effect says that we consider equivalent (or, in the present discussion comparable) those scores that correspond to equal standard score deviates. Or, if we wish, we can say that we consider comparable those scores that correspond to equal percentiles. Now, if the distributions on Form x and on Form y have the same shape, then this latter procedure will result in a straight-line conversion quite similar to the one first described, which assumes that differences in shape are of no consequence. On the other hand, if the distributions are different, then this latter conversion will be curvilinear. Still another method is one in which half the group takes Form x followed by Form y, the other half taking y followed by x. Here the assumption is made that the effect of practice is constant and unrelated to score level, and also that the effect on Form y performance (in terms of standard scores) is the same as the effect on Form x performance. Finally, the conversion between x and y is achieved by averaging the means and variances on Form x and also on Form y and setting them equal. Incidentally, the problem of practice effect is a troublesome one in the case of parallel tests. It is particularly troublesome in the case of nonparallel tests for the reason that, just as it is logically awkward to compare *performance* on tests that are not designed to measure the same thing, similarly it is logically awkward to compare *changes* on tests that are not designed to measure the same thing.

All other methods of equating that are known to be in frequent use involve the use of anchor tests in one way or another and are variations of a basic method developed by Dr. Ledyard Tucker and also by Dr. Frederic Lord – but with different sets of assumptions. In all of these methods the tests x and y are given to different groups; the function of the anchor test, which is given to both groups, is to adjust for any differences that may be observed in them. Most of these methods may also be used for deriving comparable scores between tests of different function. More recently – within the last two or three years – a method has been suggested by Dr. Lord for making use of anchor tests in estimating frequency distributions for groups on tests that were not actually taken by those groups. This method has been adopted for use with the equipercentile method of equating tests, a method which heretofore was possible only in those situations where the tests could be administered to random groups. I should also add that anchor test scores may be used in a method of equipercentile equating that has been suggested by Dr. E. F. Lindquist. In this method scores on two tests are taken to be equivalent if they are found to be equivalent to the same score on an anchor test.

If we attempt to adapt these methods of equating to the problem of nonparallel tests and consider the restriction that was discussed before, that such a conversion is applicable only to the specific conditions under which it was derived, we find that the following questions are pertinent: (1) How similar are the tests for which comparable scores are to be developed? (2) How appropriate is the group on whom the table of comparable scores is based when one considers the person or the group for whom the table is to be used? And (3) How much error can we safely tolerate in the particular use we have in mind? These questions are intentionally phrased in terms of degree, because this issue, like so many others, depends on considerations of degree. Categorical answers are inappropriate. Thus, in considering the first of these three questions, if the tests are extremely different – if, say, one is a test of quantitative skills and the other is a measure of interest in sports – then, clearly, it would be absurd to contemplate a conversion that would make any sense at all. On the other hand, if one is a verbal test heavily loaded with items of the definition type, while the other is a test of reading comprehension, then it is certainly reasonable to consider a table of comparable scores for these two tests.

With regard to the second question, if the group on whom the table is based consists of seniors majoring in literature, and the group for whom we wish to use the conversion consists of high school sophomores enrolled in a manual training course, then obviously the use is inappropriate. On the other hand, if the groups are both freshmen in liberal arts in colleges that are similarly selective, then the use is probably appropriate.

The third question, which calls for a commitment to a maximum tolerable error, is fundamental to the entire issue. Commitments regarding the amount of difference between the tests that may be allowed and commitments regarding the difference that may be allowed between the groups under consideration depend ultimately on the demands for precision that are implied by the purpose of the conversion table.

Ordinarily we are confronted with the situation where the tests and the use for which a table of comparable scores is required may not be questioned or altered. In such instances we are free to choose the kind of group we wish to use in forming a conversion table, in order to achieve a degree of error that we say we are willing to tolerate. Three such groups are: (1) the national norms group; (2) a set of differentiated norms groups; and (3) the local norms group. A fourth method – of basing a table of comparable scores on a group of people for whom data happen to be convenient and available – is simply to go through the motions of deriving a table of comparable scores without considering its meaning. The usefulness of such a table – if there is any – is extremely restricted.

Of the various kinds of comparable scores, the one based on differentiated norms groups is probably the most defensible. This procedure would yield a number of conversion tables, each based on, and appropriate for, a different norms group. If the tests are measures of the same psychological function, then the various conversion tables will, for all practical purposes, be the same. However, if the tests are measuring different functions, then there will be as many conversions as there are distinguishable and different norms groups. Each conversion, like a profile, will be descriptive of the group on which it is based and applicable only to that group. The user will be forced to choose the appropriate table with care, keeping in mind the group for which he intends to use it and the purpose for which it is to be applied.

The local norms approach to comparable scores is similar to the one involving differentiated norms, and is in general as highly recommended for the purpose of comparable scores as are local norms distributions themselves for the purpose of evaluating relative status. Here the cautions that need to be exercised are: (1) that the group has not been directly selected on either of the scores involved in the conversion; (2) that there are sufficient cases to yield reliable conversions; and (3) that they be applied only in the school where they were developed or in schools known to be similar to it.

The national norms approach is probably the least satisfactory of all, for the reason that it is not likely to be highly appropriate except when the tests in question are quite similar in function. Its primary advantage, however, is that it is the most readily applied method of obtaining rough conversion tables, if for no other reason than the fact that national norms for tests are generally readily available. The significant concern here is that none of the students on whom the comparable scores are based took both the tests in question, and that the norms groups for the different tests may not have been selected in the same fashion in order to satisfy even approximately the requirement of randomly equivalent groups that is ordinarily imposed in parallel-forms equating. And even if they were so selected, there is enough unreliability in norms samples to introduce serious errors in the tables of comparable scores. There are situations in which errors of given magnitude are tolerable in norms, but not at all tolerable in tables of comparable scores — again, of course, depending on the intended use of the tables.

I would like also to mention a method that is appropriate for developing comparable scores for situations in which candidates are permitted to take the particular test that is most appropriate to their own curricular background or interests. As applied to the achievement tests in the College Board program, the procedure is to correlate the SAT-Verbal and Mathematical scores with each achievement test, and use them in combination as a composite control or anchor test in estimating the achievement test performance of a group with certain defined characteristics, as measured by the SAT. The purpose of all this statistical work is to relate the scores on the various achievement tests to one another and to the SAT-Verbal and Mathematical scale, and to enable us to say, for example, that a 600 score on the French test is comparable, in the sense defined by the statistical method, to a 600 score on the physics test. Obviously, there are logical difficulties in expressing this comparability, as there would be with any comparability, and perhaps more than there would be with most. However, because of the way in which the achievement test scores are used, and have to be used, by the admissions officers, some kind of comparable scores must be prepared. And of the various kinds that are available, this one, which balances out the effects of differential selection for each of the groups that take the various achievement tests, has obvious advantages.

Another approach to the question of comparable scores involves the use of the ordinary correlation and prediction model. This method has at times been suggested for use with tests at the same grade level in order to provide interchangeable scores on different tests. However, as described earlier, it is clearly inappropriate here because it fails in its very purpose. One of the scores is actual and the other is regressed; they cannot be interchanged. On the other hand, when the tests are designed for different grade levels, then the method does have merit. For example, an attempt has been made to use the regression model for relating scores on the School and College Ability Test taken in the early years of secondary school to scores earned on the SAT in the twelfth grade. Since these results will be used solely as predictions – to enable students to get a preliminary idea of how well they may expect to do on a test used to select college freshmen – and not to merge or compare them with actual scores, the dangers that were described as inherent in regressed scores are not present here. Furthermore, since the presentation lays great stress on the presence of errors of estimate, it is felt that misinterpretations will be kept to a minimum.

In general, the different uses of comparable scores require different levels of precision. Where, for example, the need for comparable scores arises from a desire to use scores on different tests interchangeably for admissions purposes, the highest degree of precision is required for comparability. Otherwise, a college which gives serious consideration to test scores in admissions decisions might find itself admitting too large a proportion of applicants from one group or another because of the use of an inappropriate conversion table. On the other hand, when scores on different tests are used interchangeably for placement purposes, the demand for precision is less serious, since erroneous placements can be corrected early in the course. Also, the use of conversion tables for estimating the comparative abilities of *groups* of individuals is one that does not ordinarily require a high degree of precision. For one thing, in such comparisons the careers of individuals are seldom at stake (as would be true in admissions decisions). For another, it is a characteristic of conversions that there is relatively little error at the mean. In contrast to this, conversions of scores for individuals of high or low ability are much more likely to be in error.

It probably bears repetition that the problem of comparable scores is simply another one of many other problems about which it is impossible to make categorical evaluations that hold fast in all situations. If the decisions that are made on the basis of a table of comparable scores are made in the central region of the distribution, if the tests are relatively similar in function, if they have approximately equal reliabilities and validities, and finally, if generalizations are restricted to the kind of group that formed the basis of the conversion in the first place, these are the conditions that allow the free use of comparable scores. To demand that all these conditions hold is undoubtedly unrealistic and unnecessarily restrictive as well. At the same time, if, in any instance, too few of these conditions are operative, then testing organizations would probably do well to heed the advice that they like to give their test users, and proceed with caution.

Luncheon Address

Approaches to the Measurement of Intellectual Disposition*

T. R. McConnell,
*Center for the
Study of Higher Education,
University of California,
Berkeley*

MUCH OF THE Center's early work on student recruitment to higher education was based on the hypothesis that institutions and groups of institutions may be differentially selective or attractive, not only in ability (aptitude or achievement), but also in certain other attributes of personality. It also wished to devise means of discriminating between defined groups of students within or among institutions. The interest in differential recruitment led naturally into investigations of changes in students' characteristics over the college years in describably different institutional cultures or subcultures. In all this work, the concentration was on evidences of manifest or potential intellectuality, and, as the studies progressed, on other attributes favorable or unfavorable to the actualization of the individual's intellectual promise.

The Omnibus Personality Inventory

One of the first tasks, of course, was to select or devise measures of intellectual disposition and of certain characteristics related to intellectual functioning. Dr. J. G. Darley, with the assistance of Dr. M. P. Weissman, searched for scales either of intellectuality or emotionality which seemed to offer the most promise of distinguishing one student body from another or one sub-group of students from another.

---

* Much of the work referred to in this paper was carried out by Dr. Paul Heist, Dr. Harold Webster, and Dr. George Yonge at the Center for the Study of Higher Education.

The set of scales used in the earliest studies has since been extensively revised. Particular scales have been deleted or added, and most of the ones retained have been considerably refined. In the process, a degree of reliability was sacrificed in order to measure major characteristics as independently as possible. A detailed history of this scale development is not necessary here, since it is summarized in the *Research Manual* for the *Omnibus Personality Inventory,*[1] as the instrument has been called.

The present *Inventory* contains the following principal scales: Thinking Introversion-Extroversion, Theoretical Orientation, Estheticism, Complexity, Autonomy, Developmental Status, Impulse Expression, Schizoid Functioning, Social Introversion-Extroversion, Religious Liberalism, Social Maturity, Masculinity-Femininity, and Repression-Suppression. The original forms of these scales came from many sources. Thinking and Social Introversion-Extroversion came from the Evans-McConnell T-S-E Inventory and are little changed except in method of scoring.[2] Theoretical Orientation and Estheticism are new scales developed by the staff and validated against the corresponding sections of the Allport-Vernon-Lindzey *Study of Values.* Complexity is the outgrowth of research originally conducted by Barron of the Institute of Personality Assessment and Research.[3] The original forms of the scales of Developmental Status, Impulse Expression,[4] Social Maturity and Repression-Suppression were developed by the Sanford group at Vassar; the scale of Impulse Expression has been considerably revised. A cluster of items from the *MMPI* was the nucleus of the Schizoid Functioning scale. Religious Liberalism and Masculinity-Femininity are scales devised by the staff. Autonomy is a shorter version of the Social Maturity scale.

The KR21 reliability coefficients of the presently used scales range from .71 to .93, with ten of the thirteen .80 or above, and six .85 or above. The test-retest reliabilities on a small group of students varied from .76 to .94.

Since the staff was especially interested in intellectual disposition and intellectual functioning, it has focussed its attention on certain scales. What these scales presumably measure is summarized below.

The Inventory Scales

Thinking Introversion (TI) (60 items): Persons scoring high on this measure are characterized by a liking for reflective and abstract thought. They express interests in a variety of fields, such as literature, art, and philosophy. Their thinking tends to be less dominated by objective conditions and generally accepted ideas than that of thinking extroverts (low scorers).

Theoretical Orientation (TO) (32 items): This scale measures interest in science and in scientific activities, including a preference for using the scientific method in thinking. High scorers are generally logical, rational, and critical in their approach to problems.

Estheticism (Es) (24 items): High scorers have diverse interests in artistic matters and activities. The content of the statements extends beyond painting, sculpture, and music to literature and dramatics.

Complexity (Co) (27 items): High scorers are tolerant of ambiguities and uncertainties, are fond of novel situations and ideas, and are frequently aware of subtle variations in the environment. Most persons high on this dimension prefer to deal with complexity, as opposed to simplicity; are disposed to seek out and to enjoy diversity and ambiguity; and have an experimental attitude rather than a fixed way of viewing and organizing phenomena.

Autonomy (Au) (40 items): The characteristic measured is composed of nonauthoritarian thinking and a need for independence. High scorers are sufficiently independent of authority, as traditionally imposed through social institutions, that they oppose infringements on the rights of individuals. They are nonjudgmental, realistic, and intellectually liberal.

Impulse Expression (IE) (75 items): This scale assesses a general readiness to express impulses and to seek gratification either in conscious thought or in overt action. The high scorers value sensations, have an active imagination, and their thinking is often dominated by feelings and fantasies.

Schizoid Functioning (SF) (74 items): The high scorers admit to attitudes and behaviors that characterize socially alienated persons. Along with feelings of isolation, loneliness, and rejection, they may intentionally avoid others and experience feelings of hostility and aggression. The ego weakness of high scorers may be characterized by identity confusion, daydreaming, disorientation, feelings of impotence, and fear of loss of control.

Profile of Intellectuality

We are inclined to look on the scales of Thinking Introversion, Estheticism, Theoretical Orientation, and Complexity collectively as indicative not only of interest in ideas but also in the free play of ideas; as indicative of flexibility rather than rigidity in thinking, of the willingness to suspend judgment or search for novel solutions, of a preference for intellectual nonconformity rather than commonly held convictions, and of the ability to impose a high level and perhaps an uncommon form of integration on concrete phenomena or specific ideas.

The intercorrelations among these four scales vary from .22 between Estheticism and Theoretical Orientation to .63 between Estheticism and Thinking Introversion. The correlation between Theoretical Orientation and Thinking Introversion is also relatively high — .62. These correlations are high enough to suggest a considerable degree of agreement, yet low enough to permit some specificity of measurement. Apparently these scales tap independence as well as intellectual interest; for example, the correlation of Autonomy with Estheticism is .25, and with Theoretical Orientation, .72. Impulse Expression is correlated with Complexity to the extent of .48, and slightly with the other three scales. As the earlier research of Barron indicated, Complexity spans the intellectual, the authoritarian, and the emotional dimensions of the composite Inventory.[8]

In a sample of 372 students in four colleges under study, the correlations of the scales with SAT-V ranged from −.01 to .58, and with SAT-M, from −.01 to .43.

Loadings for unrotated and rotated (both Quartimax and Varimax) factor analyses are given in the *Manual*. In both instances, factors with heavy loadings on Thinking Introversion, Theoretical Orientation, Estheticism, and Complexity appeared, with either a moderate or a strong loading on Autonomy. In the case of the unrotated analysis, a strong component of Impulse Expression was involved. Inhibition factors emerged in both analyses. (Other factors appeared, but will not be noted here.)

The *Research Manual* gives intercorrelations of *OPI* scales with several other inventories – the Allport-Vernon-Lindzey *Study of Values*, the Kuder *Preference Record*, the *Myers-Briggs Type Indicator*, the Stern *Activities Index*, the *California Personality Inventory*, and the *Strong Vocational Interest Blank*. For the most part, the relationships seem to be what one would expect, although, of course, it is dangerous to infer comparability from scale names. As an example, correlations between certain scales of the *OPI* and of the Stern *Activities Index* may be interesting. In a group of 59 college students there were statistically significant correlations between Thinking Introversion and Change of .28, and between TI and Understanding of .57. There were correlations between TI and Play and Sex of −.40 and −.30, respectively.

Differentiation of Student Bodies

How do the *OPI* scales, especially those related to intellectual disposition, discriminate the student bodies of particular institutions or groups of institutions, as well as sub-groups within institutions? Only a few examples of differentiation can be reported here. (In these examples, the same form of some scales of the *OPI* was not always used, but the correlations between forms would be relatively high.)

Colleges with a large percentage of graduates who had become scientists and scholars attracted male and female National Merit Scholarship (NMS) students who had significantly *higher* mean scores on Complexity and *lower* mean scores on Authoritarianism than those who attended less productive institutions. The NMS women who attended the highly productive colleges also had significantly higher mean scores on Thinking Introversion. These differences were made all the more striking by the fact that the samples in the two groups of institutions were matched on combined scores of the Scholastic Aptitude Test. The greater intellectual orientation and possibly greater potential creativity of the students who went to the highly productive institutions was confirmed by the fact that they also had significantly higher mean scores on the Theoretical and Esthetic scales and significantly lower mean scores on the Religious scale of the Allport-Vernon-Lindzey *Study of Values.*[6] In a number of the Center's investigations, this combination of Theoretical, Esthetic, and Religious A-V-L scores has proved to be characteristic of students with serious intellectual interests and motivation.

The staff was also interested in differences in *OPI* scores of National Merit Scholarship students who attended institutions grouped in the following categories: Ivy League universities, other private universities, liberal arts colleges, and public universities — a rather crude institutional classification, to be sure. The students who attended the Ivy League universities differed significantly from those in the public universities on Thinking Introversion, Complexity, Originality (a scale since dropped from the Inventory), and Authoritarianism.

Differences among institutions categorized by type of control were also explored. The categories were Protestant, Catholic, private nonsectarian, and public. From each of these categories, 40 male NMS students were matched with respect to scholastic aptitude and field of specialization. No statistically significant differences on the *OPI* tests named above were found between Protestant and private nonsectarian institutions, or between public and private nonsectarian ones. There were fairly sharp differences, however, between the NMS students in Protestant and Catholic colleges. In the former, the group had significantly higher scores on Thinking Introversion and Complexity, and significantly lower scores on Authoritarianism. However, on Thinking Introversion and Complexity the differences between the students in the public and Roman Catholic institutions were not significant. These data were for freshmen.[7] Whether the same institutional differences would characterize seniors has not yet been determined.

The differences in the characteristics of freshmen between two small, highly selective liberal arts colleges may be especially interesting. The mean Scholastic Aptitude Test scores of the freshman classes of 1958 differed by approximately half a standard deviation. However, on nine *OPI* scales, including Thinking Introversion, Complexity, Theoretical Orientation, Estheticism, and Impulse Expression, the means of the two classes were about two-thirds of a standard deviation apart. The college with the lower mean aptitude score had the higher mean *OPI* scores. The group with the higher *OPI* means can be said to have a higher intellectual commitment and to be freer to pursue intellectual interests. "In the realm of ideas they may well be the risk-takers and experimenters."[8]

Differentiation of Student Sub-Groups

In the college with the lower mean *OPI* scores, students identified three major sub-groups. Group A included the socially oriented, with high school spirit, and strong emphasis on vocational goals. Group B included students who were unaffiliated with particular social groups. A majority of them were described as having a strong scholarly orientation, and many appeared to be seriously interested in public affairs. Group C was composed of intellectual liberals who tended to be critical nonconformists; many of them were majoring in the humanities and the arts.

On most *OPI* scales, the mean scores of students in Group C were a full standard deviation above those of students in Group A, and substantially higher than those of Group B. Furthermore, on Thinking Introversion and Complexity the means of Group C were higher than the freshman means in the school with the higher mean *OPI* scores.[9]

We have come to call a proportion of students in Group C "intellectual nonconformists," and Dr. Paul Heist will soon present a paper on their background and behavioral characteristics. On the basis of present information, these students may be described as being strongly inclined toward intellectual autonomy, as well as toward independence from others and from social institutions; and as having wide interests, unusual curiosity, and a critical attitude toward their environment. Many of the male nonconformists do not complete the college of first choice and many of the women never finish; those who do are likely to pursue the intellectual life generally, and to find careers in research or the creative arts. Even those who finish, however, may find adjustment difficult, for they and the colleges are likely to be basically incompatible. Dr. Heist has pointed out that some of the most academically selective liberal arts colleges and technological institutions lose more creative youth than they educate.[10]

It is illuminating to study differences between individual students as well as between student bodies or sub-groups. One such comparison between two students shows how verbal aptitude may be inversely related to some of the attributes of personality under study. The two students were two standard deviations apart on SAT-V, although close together in SAT-M. The student with the lower verbal score was high on certain *OPI* scales, and the one with the higher verbal score was correspondingly low on the same personality scales. The difference in Social Maturity was three standard deviations; in Complexity, two; in Estheticism, one and one-half; in Thinking Introversion, one; and in Impulse-Expression, one and two-thirds, in reverse order to the difference in verbal level. The student with the higher verbal aptitude was immature and uninterested in intellectual pursuits, and had an ordinary level of readiness to express impulses. The one with the lower verbal aptitude, in contrast, was much more mature, had a high but not undisciplined energy level, and possessed a serious interest in intellectual pursuits. The two students probably would respond quite differently to a common educational environment.[11]

In exploring the personality characteristics of known student groups, staff members compared the attributes of students in the division of Creative Arts at San Francisco State College with those of a large sample of students at the University of California, Berkeley. Since there were differences between men and women, the investigators reported their scores separately; they also provided data, not only on the entire group, but also on those majoring in art, music, and drama. Only a portion of the data is summarized below.

The Creative Arts group was higher than Berkeley freshmen in Thinking Introversion, Complexity, Estheticism, and Nonauthoritarianism (Autonomy), but not in Theoretical Orientation. This general pattern held for drama, art, and music majors separately, although the women invariably had higher scores than the men on Thinking Introversion, Estheticism, and Complexity.

The mean of a small group of students (N = 10) rated high on creativity by faculty members was approximately half a standard deviation higher than the entire group of creative arts majors on Thinking Introversion, Estheticism, and Complexity, as well as on Theoretical Orientation.

In studying the careers of students in the three highly selective liberal arts colleges, the staff has become interested in those who fall by the wayside, or who transfer to other institutions. Heist and Webster have computed mean score profiles on two groups of students who dropped out of the college referred to above as having the highest mean scores on such *OPI* scales as Thinking Introversion, Estheticism, and Complexity. Out of one group of five women identified by three deans as exceptionally bright nonconformists, only one graduated, in spite of the fact that the group's mean profile was from one-and-one-half to two-and-one-half standard deviations above the *OPI* norms, and as much as one standard deviation above the institutional mean on Thinking Introversion, Complexity, and Estheticism. With such a pattern of characteristics, these students should have been more than able to meet the intellectual demands of this or any other college. However, their mean Impulse Expression score was nearly three standard deviations above the norm, and about two standard deviations above the mean of all freshmen in the particular college. Furthermore, their score on the Schizoid Functioning scale was one standard deviation above the norm and about two-thirds of a standard deviation above the college mean. These students apparently were not sufficiently in control of their impulses and not sufficiently stable emotionally to meet the educational demands they confronted.

As mentioned above, the Center is interested in the actualization of potential intellectuality. There is reason to believe that the scales which probe impulsivity, alienation, and emotionality will assist us in distinguishing intellectually oriented students who make constructive use of their talents from those who fail to do so.

The second group who dropped out of the college in question were men who had Thinking Introversion, Estheticism, and Complexity scores which were about at the mean of the norm group but almost a standard deviation below the college freshman mean. Their Impulse Expression score was a standard deviation above the norm, but about at the particular college mean, and their Schizoid Functioning score was approximately at the norm and near the college mean. These students appeared to be sufficiently disciplined and emotionally stable to succeed where they were, but their intellectual interests were incongruent with those of the student body as a whole and with the intellectual press of the institution. Eleven of the fourteen had dropped out by the beginning of the third year.

Theoretical-Applied Dimension

I should like to describe one more means of measuring intellectual disposition. Dr. M. P. Weissman, counseling psychologist of the Student Counseling Center at Berkeley, has devised a novel method of distinguishing students with theoretical from those with applied intellectual dispositions.[12] By theoretical disposition, he meant "a preference for the abstract and speculative, and the satisfactions to be gained from the interplay of ideas. Individuals in whom such preferences are strong might be expected to gain satisfactions from intellectual activities intrinsic to problem-solving tasks, or to enjoy learning for its own sake. The activities are ends in themselves and the intellectual work may be regarded as being 'functionally autonomous'."

[illegible] applied disposition, he meant a preference for the utilitarian and practical, and assumed that in this case intellectual activities, even creative ones, would be means of attaining utilitarian goals.

He hypothesized that either a theoretical or an applied disposition might cut across many occupations and operate at various vocational levels.

Weissman devised a method of pattern analysis for the *Strong Vocational Interest Blank* which permitted him – and other qualified judges – to combine a preference for a subject or occupational field (e.g., science or business) with a theoretical or applied attitude toward that concrete interest. The method of classification involves the intensity of a core group of keys in the individual's profile, the strength of the subject area, and certain related characteristics. The core group, which is the crux of the classification scheme, is composed of the scales for artist, psychologist, architect, physician, and author. If a profile represents a theoretical disposition, three of these scales and either artist or author must have a score of B-plus or above. There are three sub-classifications of theoretical disposition. The first is designated as "abstract"; in this instance, intellectual curiosity centers on the world of ideas, concepts, and theory, often without tangible subject or occupational focus. The second sub-category is theoretical-scientific; the emphasis is on liking for speculative and creative thought. the spirit of inquiry, and the scientific method; intellectual curiosity centers on science and research. In the third subgroup, intellectual interests center on social institutions, customs, and behaviors.

The second basic category, between the most theoretical and the most applied, is applied-professional. Here the core group is insignificant, but certain subjects or occupations become necessary components. Five sub-categories appear. labelled biological science, technical, welfare, business, and verbal. The orientation in each case is to application, but also to independent responsibility. Level of operation is important.

The final main group is called applied-technical. Here the orientation is again applied, but aspirations for responsibility are limited.

After classifying some 900 National Merit Scholarship freshmen into these three broad categories, Weissman studied their preferences for fields of specialization, occupations, and institutions of higher learning. He also explored differences among the three groups on certain scales of the *Omnibus Personality Inventory*, especially the scales for Thinking Introversion, Complexity, Originality, Authoritarianism, Impulse Expression, Social Introversion, and certain scales from the *MMPI*. The detailed differences cannot be reported here, but the composite pictures Weissman drew from the data are summarized below.

Group A

The individual tends to be thoughtful, reflective, independent, creative. He responds to complexities of the environment and internal impulses. He is sensitive to others, permissive in relationships, and tolerant in attitude though somewhat lacking in determination and persistence.

He chooses a major in the natural sciences, humanities or social sciences approximately twice as often as in the professions or business; and in the natural sciences two and one-half times as often as in engineering. He chooses the humanities or the professions relatively infrequently, but prefers the former and prefers it more frequently than students in either group B or C.

He tends to become a physicist, chemist, mathematician, research scientist, or college professor. He selects a vocation in the arts or humanities infrequently, but is more than 15 times more apt to do so than to choose one in business or politics. He favors institutions in the top twenty most productive of scientists and scholars over all others in the ratio of 1.6 to 1.

Group B

The individual tends to have only a moderate liking for abstractions and reflective thought. He is outgoing, socially participative, aspiring, responsible, serious, steady, conforming, self-controlled.

He chooses preprofessional and business majors twice as often as those in the humanities and the natural and social sciences. He prefers engineering one and one-half times as often as physics, chemistry, or mathematics. He is likely to become an engineer, physician, lawyer, teacher, accountant, or business man. He favors institutions of low productivity over the top twenty in the ratio of 1.4 to 1.

Group C

The individual tends to prefer physical activity to thoughtful reflection, to deal with the simple and concrete rather than with the complex or intangible, and to desire power, authority, and directive relations with people. He tends to be socially aloof.

He chooses majors almost exclusively in technical fields, electing engineering three times as often as the natural sciences. He seldom selects a major in the humanities, the social sciences, business, or politics. He favors institutions in the "low" productive category to the high twenty in the ratio of 2.2 to 1.

Both at the Center and elsewhere, the Weissman technique of classification has been used to differentiate known groups of students. For example, Heist and Williams studied the correlates of different levels of achievement among freshmen at the California Institute of Technology. This group of students is highly homogeneous in scholastic aptitude (mean SAT scores: V, 649; M, 734). On the basis of freshman grade point average, the investigators selected three groups of high, average, and low achievers (N = 37 in each group) whose average verbal and mathematical scores on the SAT were not significantly different. They found that the groups differed significantly on the Weissman SVIB profiles. Almost three times as many high as low achievers displayed a theoretical orientation. On the other hand, twice as many low achievers and half again as many average achievers as high achievers appeared in the applied-professional group.[13] The differences in intellectual disposition between these two groups are not especially surprising, since there is evidence that the academic program at the Institute has a strong theoretical rather than a pronounced applied bias.

Studies of Changes in Personality Characteristics

As noted at the beginning, the Center's studies of differential student characteristics led to investigations of student development. The staff is now analyzing changes in students which occurred during the college years in eight institutions. The Center is also far along in the assessment of changes in a group of National Merit Scholarship students. In both instances, the emphasis is especially, although not exclusively, on students' intellectual interests and dispositions at the time they entered, and what happened to these characteristics while they were in college. Course grades say something about a student's intellectual accomplishments, but they do not say enough. For this reason, changes in attitudes and values, and in the characteristics measured by the *OPI* scales of Thinking Introversion, Theoretical Orientation, Estheticism, and Complexity are matters of central concern.

It is clear from the analyses already made that significant changes occur on these and other scales. A report of mean changes in the National Merit Scholarship group on the *OPI* and other instruments will soon be available. In addition to mean changes, many individual changes in both directions have been found, and these will be presented in due course.

Concurrently with these studies, other longitudinal investigations have been under way. One is of students at the California Institute of Technology. Another is of comparable groups of high school graduates in sixteen cities who did and did not go to college.

It is to be hoped that out of this wide range of studies, a more explicit account of the intellectual development of college-age youth can be written.

REFERENCES

1. Paul Heist, T. R. McConnell, Harold Webster, and George Yonge, *Omnibus Personality Inventory – Research Manual.* Berkeley, Center for the Study of Higher Education, 1962.
2. Catharine Evans and T. R. McConnell, "A New Measure of Introversion-Extroversion." *Journal of Psychology* 12:111-124, 1941.
3. Frank Barron, "Complexity-Simplicity as a Personality Dimension." *Journal of Abnormal and Social Psychology* 48:163-172, April 1953.
4. Nevitt Sanford, Harold Webster, and Mervin Freedman, "Impulse Expression as a Variable of Personality." *Psychological Monographs* Vol. 71, No. 11, 1957.
5. Paul Heist, T. R. McConnell, Harold Webster, and George Yonge, op. cit.
6. Paul Heist, T. R. McConnell, Frank Matsler, and Phoebe Williams, "Personality and Scholarship." *Science* 133:362-367, February 10, 1961.
7. E. D. Farwell, J. R. Warren, and T. R. McConnell, "Student Personality Characteristics Associated with Groups of Colleges and Fields of Study." *College and University* 37:229-241, Spring 1962.

8. Paul Heist and Harold Webster, "Differential Characteristics of Student Bodies — Implications for Selection and Study of Undergraduates." In T. R. McConnell (editor), *Selection and Educational Differentiation*. Berkeley, Field Service Center and Center for the Study of Higher Education, University of California, 1960, pp. 91-106.
9. Paul Heist and Harold Webster, op. cit.
10. Paul Heist, "Reflections on the College Education of High Ability Youth." Berkeley, Center for the Study of Higher Education, University of California (Mimeographed).
11. Harold Webster, M. A. Trow, and T. R. McConnell, "Individual Differences Among College Freshmen." Chapter IX in H. B. Nelson (editor), *Individualizing Instruction*, Sixty-First Yearbook, Part I, of the National Society for the Study of Higher Education. Chicago: University of Chicago Press, 1962.
12. M. P. Weissman, *An Approach to the Assessment of Intellectual Disposition Among Selected High Ability Students*. Ph.D. dissertation, University of California, Berkeley, 1959.
13. P. A. Heist and P. A. Williams, "Variation in Achievement within a Select and Homogeneous Student Body." *Journal of College Student Personnel*, Vol. III, No. 2, pp. 50-59, December 1961.

Session III

Theme: Processing Educational Data

Information Systems in the Educational Environment

CHARLES R. DECARLO,
International Business
Machines Corporation

BEFORE PURSUING the application of information technology to the field of education, it would be well to consider the need for such application in terms of the people affected. The development and application of information-handling systems, particularly when coupled with testing and selection, are worse than wasteful if not done for purposes which fundamentally improve the educational process and create greater opportunities for more individuals.

In spite of the several books which have recently appeared, viewing with alarm the abuses of testing, it seems clear that the establishment of standards, and methods for testing to those standards, far from denying the unusual genius, has in fact broadened in a democratic way the availability of our finest educational resources. Even in those cases where the "missed genius" and "late bloomer" may not be recognized by testing procedures, the undeniable upgrading of our educational process through the agency of curriculum development, testing, emphasis on excellence, etc., now provides a better environment for their particular talents to shine at a time appropriate to their natures.

Illustrating this point are the following remarks from Dr. L. A. du Bridge, President of California Institute of Technology:

"There is no way on earth we could have found a group of students in the top 1 or 2 or 3 or 4 per cent of the entire country if we had not had something like the College Board tests to go on. The experience at Harvard, Yale, Princeton, MIT and Stanford is similar — a rise of from 10 to as much as 40 points on the percentile scale for the average freshman has occurred in the past ten years. Princeton reports that their entire list of applicants last year was above the level of those actually admitted ten years ago.

"The old-style admissions procedure for some old-line universities was to depend on graduates of a few fine old-line private schools, ignoring a host of graduates of public schools because there was no adequate yardstick for comparison. Today the typical university draws students from high schools of all types all over the country. Our 182 freshmen last year came from 173 high schools, only 25 of which were private schools. And nearly half of our freshmen came from schools east of the Mississippi River. MIT reports that its 900 freshmen came from 615 high schools — 87 per cent of them public schools. And 87 per cent of their students came from outside of New England.

"And have these students performed in college in a way to bear out their high test scores? They have indeed. We find very close correlation between college performance and entrance test scores. And, as the freshmen have improved, we have had to advance the quality and content of our college courses in order to keep up with them. Most of our freshmen of ten years ago would have flunked dismally in competition with our freshmen of today — except, of course, if the freshman of ten years ago could have gone to the high school of today, he would have done better also.

"Thus, on the basis of both test scores and college performance, we can say that the freshmen at these universities are on the average brighter than they were ten years ago. That, of course, is because we can be more selective than we used to be."

Another aspect of the need for the application of new technology in the educational process is our urgent demand for education for more people, in greater depth and breadth, and with higher efficiency than ever before. It is estimated in the next decade that our high school enrollment will increase by 50 per cent while higher education will increase 100 per cent. The pressure of this fact on educational planning will insist that methods be developed to assist our dedicated teaching profession in accomplishing their tasks. Obviously relief from many of the administrative, clerical, and routine tasks, now a part of the teacher's work day, will help. Also, new developments in the teaching experience itself should lead to more efficiency in terms of cost and use of facilities. At the same time teacher-pupil relationships should be made more meaningful.

Another ingredient which must be added in the new education is the development of a firm commitment on the part of all to a life of continuing educational process. The world we live in, as well as that which lies ahead, suggests that our true vitality will be measured in our ability to manage change, particularly as it affects the individual, our values, and our institutions. Professor W. L. Warner characterizes this American phenomenon well when he says:

"I believe that the nature of our collectivity is such that change is built into it and is an essential part of it; that to be what it is at any one moment in time, this society must continually change and become something else. The processes of change are such that, while the forms of the past are being absorbed into the present and losing their identities, the present ones, re-formed by the future, are becoming something different from what they now are.

"In such an emergent process, the past influences the present; and the processes that timewise become the events of the future are already present and, operating there, make their significance felt today before tomorrow's world takes over. Quite literally these activities go on now in the bodies of each of us."

Implicit in this statement is the demand that we have a continuing awareness of the nature of our present world, as well as the forces leading us to new environments. The attitudes and abilities of a people capable of meeting such a future must become, in the general populace, the same as those that are found in our most capable people today. I believe it is possible to "move up" the population through a strong commitment to education. In the field of adult education, vocational training, and industrial education and training, the uses of standards, testing and information-handling techniques will help our leadership in discharging this commitment.

Finally, a word must be said for organizing the uses of information technology in the field of information retrieval, transmission, and display. These fields bear directly upon the ability to organize new curricula based upon structure, as well as fact.

Typical of the past, in Dickens' *Hard Times* Mr. Gradgrind observes:

"Teach these boys and girls nothing but Facts. Facts alone are wanted in life. Plant nothing else, and root out everything else. You can only form the minds of reasoning animals upon Facts; nothing else will ever be of any service to them."

However, in today's world the proliferation of fact, of knowledge, and of experience, eliminate the possibility of one person being truly, and completely competent in many disciplines. The modern ideal equivalent to the Renaissance man, it would seem, is a person who has a deep appreciation for the unity and structure of knowledge, who has developed particular competence in several fields of knowledge, and who possesses a continuing curiosity about, and concern for, the world in which he finds himself. Such a person's accomplishment within his chosen fields are very likely to be outstanding. In addition to this he will be a contributor and active participant in the several worlds which make his life.

All of the above remarks are background against which speculations and plans for the future in educational technology should be measured. However, it is instructive to examine first the range of functions machines can perform today in assisting the educational process:

1. There exists a host of data processing applications in the administrative and statistical work of the educational plant. Performing these applications, about 1,000 school districts today have access to data processing equipment of some kind. Between 500-750 have their own installations and by the end of this year approximately 100 will have high speed computers. Thus such activities as census and attendance records, registration, appropriation accounting, payroll, etc., are sus-

ceptible of application today—subject only to the normal constraints of cost, size, etc.

2. There are available test-scoring machines which can read either cards or papers containing discrete marks. These, combined with the development of multiple choice tests, have created a whole range of tools for the use of teachers, counsellors, administrators, and educational researchers.

3. There are today some applications in programmed learning in which electronic and electrical technology are being used. These are in addition to the new type of programmed text book. Most of the computer system work done thus far has been of an experimental nature but yields great promise for use.

Specific examples of this type of work are described by W. R. Uttal in *IBM Research*. In his work he has used computing systems to teach stenotypy, scientific German, and psychological statistics. In each of these courses the machine system is made adaptive to the student's learning rate. The total relationship is referred to by Mr. Uttal as Conversational Interaction. Measures of retention, efficiency, etc., are made in published reports.

4. Finally, the use of computers in the universities has developed to the point where several hundred institutions now have systems installed. Work carried out on these systems, while heavily oriented to engineering and science, has begun to broaden out to include medicine, law, social sciences, and scholarship in the humanities. Present use of these systems indicates clearly that in the future the complete range of knowledge and scholarship will be assisted by these useful tools. Indeed, in the long run the information-handling centers will be functions as important to the university as the libraries.

Looking forward to some possible applications of information technology, we can consider the application of information retrieval. Our whole history as a civilization is called "recorded civilization." Prior to recorded civilization, memory of the past must have been that which was transmitted by human communication in tribal fashion. Civilized man has attempted to keep track of and to retain his past through the use of written language, symbols, and myths. In the past several hundred years, with the advent of books and the proliferation of science, there has been generated an enormous amount of information. Now our past lies behind us with information growing at an exponential rate. The development of systems using information technology will change the way in which the past is brought to bear upon the present and the future.

Concomitant with the development of a new method of handling past information will be changes in the teaching experience. For example, we can imagine a system with enormous memory storage (comparable to a personal library, complete with research librarian) available to each student. Such a system would provide relevant material to the depth required for study in a specific field. In addition, there might be a method of course instruction determined by people with the highest competence in each field and made available through the agency of an information-processing system to the student. With this system, through the use of pencils, keys, television screens, etc., the student could proceed at his own pace, following the series and sets of instructions. In the future it should be possible to have individual displays, very much like the television screen, on which the student can see dynamic examples of the subject matter as he proceeds with the learning experience. In such an environment the teacher's contribution would be enormously amplified. The teacher would be enabled to perform the highest functions of teaching which are:

1. acting as an interpreter for what is learned – not so much concerned with the fact and content but putting it in perspective with respect to the internal structure of the subject and with respect to other disciplines around the particular subject;

2. developing in the student the sense of responsibility for his own education and participation in the learning experience;

3. differentiating among the students those that need special help, remedial help, and other counseling. In this way we will move towards the ideal of individual tutorial teaching.

Indeed, the educational environment of the future might consist of individual learning rooms in which the student has available complete and dynamic displays with which he can communicate directly by "light guns." This would enable him to have direct communication with the central information-handling system which will be presenting the course content at a pace dependent upon his past progress, retention, and current response. In addition to this, the system could provide the student audio information matching the visual presentation. This represents, in a sense, the ultimate synthesis of the language laboratory, the programmed text, and the application of new psychological and curriculum developments.

As part of the information-handling system, the learning history for each individual could be kept automatically and made available to both student and teacher to indicate areas of weakness requiring remedial measures, achievement, testing results, and patterns for future independent study on the part of the individual. Further, it would be possible for the teachers, or teams of teachers, to have contact with individuals or small groups scheduled at human convenience. The use of long distance television, cheaper video tape recording, standard methods of abstraction and retrieval, and other computer developments will literally present to the teacher and the learner an unbelievable opportunity to probe and proceed with interests in particular fields. Thus technology today, plus what seems certain in the offing, gives promise of creating a totally new environment for the learning process within the next 25 to 50 years.

Any such activities as data recording, record formats, education displays, etc., will be aided by the new electronic technology. National and international centers for the handling of data processing and curriculum needs within education will probably be developed. The recent proposal for a national education data processing center is a first step in providing the leadership along these lines. An important consideration lies in the fact that the introduction of information-handling technology into the educational environment is at a sufficiently early stage that meaningful and helpful standardization can take place now before major coordinating programs are under-taken. The direct participation of leaders in all aspects of education is needed now to prepare for the future design of educational systems. At this point it is safe to say that technological ability is not the limiting factor in the full development of the new education. Rather our limitations are the availability of sufficient numbers of creative people to understand and apply such systems, and the development of methods enabling our present institutional framework to move ahead in making the necessary changes.

With respect, however, to technological limitations, I would like to comment upon the several major barriers to future design.

At this time it seems not unreasonable to assert that education processes which involve information, and explicit, logical procedures are amenable to implementation in this new technology. Indeed, there are some enthusiasts in our field who believe most of what is done in the world can be automated — though they seldom question to what purpose or end. In fact, however, this view is totally unwarranted — for at this time there are major technological and intellectual barriers to be overcome before the next major advance occurs.

First, there is the question of physical size and the number of electrical connections required in building a machine. In the modern system, individual elements switch their state, that is, they go "on" or "off" in billionths of seconds. This is near the speed at which the electromagnetic waves themselves move. Thus, in designing machines, the flow of current cannot be thought of as instantaneous with respect to the other elements of the system. This makes the design problem analogous to a design problem in hydraulics, rather than electronics. As part of the effort to overcome this problem, the size of the system is reduced in order that electromagnetic waves travel shorter distances. This will enable machines to operate faster for a given set of elements. However, as more elements are placed into a small space for this reason, there is an increase in the difficulty of providing the number of physical wires needed to carry the information into the given space. As a result of this size problem, major research is being carried out to eliminate as much as possible wiring connections and separate elements in machines. Printed circuits and molecular electronics are part of this thrust.

A second barrier concerns what the scientist calls "noise" in the system. The electromagnetic waves used in the computer are of high frequency and very low energy. At the same time there is a certain amount of background thermal noise or "hum." This is due to the agitated movements at the atomic level within the material making up transistors, tubes, etc. So the problem is analogous to receiving a good radio signal when the signal is weak and there is much static. Research is going on to develop low temperature devices to reduce the "agitation" and minimize background noise.

The third major barrier concerns the way in which the elements of the system are organized, or put together. Today the largest, most powerful computer built is still very primitive: It still does one thing at a time, laboriously plowing through each step before going on to the next. This sequential mode of operation is executed by elements that operate in billionths of seconds so that the total effect is one of incredible speed. However, such systems have great difficulty in controlling events that must happen simultaneously in processing.

Such sequential operation is under the direct control of a centralized timing organ. Comparing this with the human nervous system, we are impressed by the fact that our bodies manage a great deal of thought and motor control with elements that operate substantially more slowly than those in the computers — slower by a factor of thousands. The reason for this is that we have evolved and developed a scheme for decentralizing, in our bodies and nervous systems, the purpose and control of our work. But we have not learned how to do this with even the most elegant machines. The system designer must organize individual parts of the machine, each working separately, into a total system. Our arms, legs, and eyes all operate somewhat independently and yet connect us through our brain as an integrated system. But we are a long way from understanding the nature of the human organization as it might be applied to machine systems. Problems of task decentralization and organizing for simultaneity in the computer are the major intellectual challenges in our field.

In conclusion, these fundamental ideas concerning the nature of information, the ability to execute explicit procedures automatically, and the ability to design reliable systems are the heart of the information revolution. This revolution is being implemented in technologies which are the result of development in the classical fields of physics and engineering. However, when the technology and the ideas outlined above are combined, one arrives at a new degree of complexity wherein a system appears to have almost human aspects.

The reason for this, of course, is that the system begins to perform work in an area that once was considered exclusively "human." Thus, the adaptation of these machines to clerical, administrative, and certain technical occupations presents us with a more subtle problem of change than the prior technical advances in "prime mover" technology. In the past an individual could compare his muscle with a steam shovel, and see the differences and similarities. Thus, in time it came to be accepted that machine power was an amplification of muscle power.

However, our inability to know how we think makes it very difficult to find the similarities and differences between ourselves and the system. Thus we have the situation in which human characteristics are assigned to "giant brains." We must bear in mind, however, that machines must have explicit and formal procedures in order to operate. This is in contrast to the human being who, through some indefinable ability, is able to recognize patterns in the most simple activities which defy our power of analysis in the computer field.

To date, only the human being is able to give purpose and meaning to his intellectual endeavors. I have yet to see the machine or system which can generate meaningful work. To generate purpose has been far beyond anything we can do today. We can amplify man's environment, we can amplify the control of the environment, but we cannot develop machines which will generate purpose or meaning in that environment.

The application of such technology in the educational environment with its great potential to better human life presents at once an opportunity and a challenge as great, or greater, than any faced by this community to date. The determination of our future in many ways depends directly upon the leadership exercised by those currently dedicating their lives to education.

The Role of Machines in Educational Decision-Making

Ralph W. Tyler,
Center for
Advanced Study in the
Behavioral Sciences

THE PRESENTATION just made by Dr. DeCarlo has provided an exciting prospect of new technology becoming available to schools and colleges. Since I have had to prepare my paper before hearing Dr. DeCarlo, I have not had the benefit of a clear and detailed picture of the developments taking place in the structure and functioning of machines. My comments are based upon my limited knowledge of the technology currently available, namely, electronic scanners, high-speed computers, large capacity data-storage units, and consoles permitting selective programs. The potential which I foresee lies in the great speed and flexibility of electronic scanners which provide varied types of input to the computer at a fabulous rate; in the speed and accuracy of the computers in making computations and comparisons; in the mammoth memories of the large capacity data-storage units; and in the increased efficiency of the machines made possible by multiple consoles which are synchronized so as to enable the large computer to work out several programs almost simultaneously. These consoles also permit output in the form of numbers, letters, graphs, printed words, and oral material. In the discussion which follows I have in mind the use of machines of this nature.

A Model of the Decision-Making Process

The other term in the topic assigned to me is "educational decision-making." It is obvious to all of us that teachers, school administrators, parents, pupils, and the public generally are all confronted with choices as they deal with education and they are consequently making decisions. In order to see more clearly how machines can be used in the process of decision-making it is helpful to consider a generalized model of major elements involved in the decision-making process. Decisions are made when action is necessary and there are two or more alternate courses of action possible. That decision is most intelligent which chooses the course of action most likely to achieve the goals sought at the least cost. The elaboration of this model in general terms indicates, first, that in most situations not one goal but several are sought, and we also recognize that in most cases each goal is achieved to a certain degree rather than simply achieving the goal or failing to achieve it. The model of decision-making will then include several goals, each achieved to a different degree, and the wiser choice will have resulted in the maximum achievement of these goals; that is, the combination of the several achievements will be greater than the combination achieved by any other choice that could have been made. Correspondingly, the cost of the course of action chosen is generally a multiple cost; that is, there will be costs of time of students, time of faculty, of money, and the like. The model will then include several costs and the wiser choice will have resulted in the maximum achievement of goals with minimum cost.

The model needs still further elaboration. Many of the situations where choices are important involve courses of action, the costs and the results of which can only be estimated with some degree of error. It is generally not possible to foresee precisely to what degree the goals will be achieved by each course of action nor to compute precisely what the several costs will be. High speed computers can provide estimates from the data of past experience using a considerable number of variables which appear to be relevant and providing a practicable multi-variable solution which could not have been obtained because of time and cost by previous methods of computation. This is an important advantage in the use of machines. Nevertheless, the data from past experience frequently differ in some respects from the data currently obtained because of changes in the situations, the human beings involved, and their interrelations. For example, some of the alternate choices may never have been taken in the past so that their consequences are not known. For these reasons, a rational strategy of decision-making will provide for a periodic review of the results and the costs of decisions made and a comparison of these results and costs with those that were estimated when the course of action was chosen. Discrepancies between the estimates and the actual costs will then be studied for clues to explain the bases for the errors in the estimates. In this way, succeeding estimates are likely to be less subject to error, and decisions more wisely made from them.

In making educational decisions, this general model of the intellectual process involved is useful. However, goals, costs, and criteria helpful in estimations of them are specific to different educational problems. Hence, I have chosen several kinds of educational problems for examination to identify the role of machines in deciding courses of action relevant to these problems.

Allocating and Scheduling Learning Resources

As schools and colleges have become larger, and as teaching personnel and facilities have become more varied, the problem of allocating and scheduling them efficiently has become more difficult. How many teachers of each subject and who have certain special skills will be needed, how many lecturers, how many counselors? How can they be scheduled so as to provide most efficiently for student learning? Similar questions must be answered for facilities – large rooms, small conference rooms, individual booths, laboratories, etc. These questions are also faced in the purchase and scheduling of equipment. The difficulty of making good estimates in answering these questions in the past has been largely due to laborious and time-consuming efforts required to work out allocations and time schedules. Usually, no such laborious computations have been made and as a result student programs and learning opportunities have often been limited because of the inefficient use of these educational resources.

Now, however, high speed computers enable us to prepare individual student programs which specify classes, group work, individual booths, laboratories, and special equipment required by each individual. When these individual cards are fed into the computer, it produces from the aggregate of the individual cards the total resources needed. Then a schedule program is placed in the computer and the individual student's schedule is printed out together with the class lists and other data needed by the teachers and counselors. Although this is not a complex problem in machine terms, its efficient solution is a helpful contribution to decision-making in the allocation and scheduling of learning resources.

Admission to College

One of the perennial educational problems is that of college admission. Some colleges have for years had more applicants than they were prepared to admit. Their admission problem has been to select from the applicants those most likely to make best use of the opportunities provided by the college. Almost all colleges have had among their applicants some who were unable to maintain the required academic standing or for other reasons were dropping out. The admission problem in these cases has been to identify those applicants who have very little chance of academic success. For both of these kinds of problems, a traditional solution has been to develop from an applicant's high school marks and his scores on entrance examinations a prediction of his college grade-point average. For colleges having to select the best candidates from a larger number of good students, the assumption that those who had the higher predicted grade-point averages were the best candidates to admit seemed to be a reasonable one. For colleges seeking to weed out candidates likely to drop out for academic reasons, the assumption that the candidates to be denied admission or advised not to enroll were those whose predicted grade-point averages were below the required standard also seemed to be a reasonable one.

Today, the situation has changed. Our economy is increasingly based on science and technology. Automation is rapidly replacing unskilled and semiskilled labor, including clerical workers as well as those in trades and industry. There are very few opportunities for employment of those who have not graduated from high school and decreasing openings for those who have not been to college. As employment in agriculture, unskilled labor and semiskilled labor drops, the demand for scientists, engineers, administrators, teachers, social workers, nurses, technicians in industry, laboratories, and health centers increases. The problem is no longer the admission to college of the few who can complete their academic work and be employed in the professions but how to identify the potential for further development in about half of all our youth and then to find the educational situations in which they can be stimulated and aided to develop into active learners growing into mature responsibilities in their roles as citizens, parents, and members of demanding occupations. The solution to this problem is not reached simply by admitting those applicants whose predicted grade-point averages are above the required minimum in a given college, and advising the others not to enroll. The problem is more complex, and now that we have new machines we have real prospects of devising better admission procedures.

Let us consider the problem first from the standpoint of the individual high school student. The most important relevant questions to which he and his parents need answers are: "In which colleges am I likely to make most educational progress if I continue with my present habits and attitudes? What changes in my habits and attitudes would be likely to result in marked increases in my educational achievement in each of the colleges under consideration?" Answers to these questions would give the student and his parent information on which to base decisions about whether to go to college, where to go to college, and what new patterns of behavior will be required to derive maximum benefit from the college experience.

At present, answers to the first question are very crude estimates. The most common estimate is the student's relative rank among high school seniors on a combination of his Scholastic Aptitude Test score and his average high school mark. This can be interpreted as an estimate of his probable academic success in a typical college, but since many colleges are not typical, it is not a very good indication for many colleges and particularly for students whose relative rank is in the middle third of high school seniors. The college grade-point average is a very inadequate estimate of how much a student has learned and unless specific estimates are obtained for each of the colleges under consideration, no answer is provided to the question "In which college will the student be likely to gain most educationally?"

The availability of high speed computer service makes it practicable to provide better estimates to this question by including more relevant information about the student, the high school, and several kinds of learning outcomes for various colleges in the prediction equations. For example, it is now feasible to include his grades and test scores in the several high school subjects, to include information on interests, study habits, peer-group memberships, and the like. It is possible to include data about the high school, such as its academic press, peer-group orientations, its relative marking system. It is feasible to develop specific predictions for different colleges using not only grade-point averages but also more objective measures such as scores on the Graduate Record Examination, and predictions in each of the major subject fields. This kind of aid to the student and his parents has only been dreamed of in the past. The new machines are bringing it into the realm of feasibility.

But, in some respects, more important for the guidance of the individual is the possibility of getting estimates on the second question: "What changes in my habits and attitudes would be likely to result in marked increases in my educational achievement in each of the colleges under consideration?" Our present estimates based on average high school marks and entrance test scores do not give the student any basis for improving his approach to learning or his habits of work. But, if we include in our prediction equations such data as the student's habits, skills, peer-group membership, we are dealing with conditions more nearly under the control of the student and his parents. Hence, the development of alternate predictions of the amount of learning in specific colleges when specified changes are made in these conditions can have practical value in suggesting courses of action for the student which are likely to increase his educational effectiveness.

Let us now consider the admission problem from the standpoint of the individual college. The most important relevant questions to which the college needs answers are: "What aggregations of students are likely to make most educational progress if the college continues with its present faculty facilities and practices? What changes in conditions in the college are likely to increase student learning and for what aggregations of students?" These questions imply two assumptions which are not commonly expressed in admission procedures. The first is that the amount students learn is affected by the composition of the student body so that the problem of selection must consider the aggregate student body as well as each student. Some evidence from studies like those of the National Merit Scholarship Corporation as well as the testimony of college teachers furnish a basis for believing in this. A second assumption is that some of the conditions in a college affect the amount of learning for some aggregates of students. Changing these conditions may make it possible to admit and to graduate some young people who are not likely to be successful academically under the present conditions.

The nearest attempt to answer the first question systematically in present practice is to select for admission those applicants with the highest relative ranks on a combination of average high school marks and entrance test scores, making some adjustment in terms of other factors deemed significant, such as athletic promise and proportions of men and women, and of in-state and out-of-state residence. We know that the college grade-point average is not a satisfactory measure of how much the student has learned. This can better be expressed as a difference function between more objective measures of college achievement and of high school achievement. With the use of high speed computers, estimates of college learning can be derived which are more satisfactory than those now in use. We also have some evidence to indicate that there are peer-group influences upon academic learning, so that estimates of how much students are likely to learn in college will need to include some measures of the characteristics of the aggregate group or groups. Data about student "needs," about the particular "press" of specified peer-groups, about the range and distribution of abilities and skills can now be included in prediction equations. For the first time it becomes feasible to make systematic estimates of the characteristics of student bodies as related to amounts of academic learning. This will provide a basis for selecting student bodies with maximum learning as the criterion. In many cases the selection will be quite different from that practiced previously.

But, in some respects, more important for society generally is the possibility of getting estimates on the third question: "What changes in conditions in the college are likely to increase student learning and for what aggregations of students?" Our present estimates of which students will succeed academically in the college give the college faculty and administration no basis for improving its educational effectiveness or for establishing conditions which would aid in the education of some groups of students who are not likely now to succeed. It is now feasible to put into prediction equations such data as: types of college "press," college peer-group "press," types of honors courses, extent of independent work. It is then possible to estimate the effect on student achievement of specified variations in these conditions. These are conditions which are to some extent under the control of the college. Hence, the development of alternate predictions of the amount of learning of specific aggregates of students when specified changes are made in these conditions can have practical value in suggesting courses of action for the college to take in order to improve its environment for the education of the aggregates of students which it wishes to serve.

Although initial efforts to develop much more adequate information for decision-making in the area of college admission will be difficult and the results will only approximate the level of precision desired, the capacity of the new machines to handle masses of data rapidly and economically is reason for optimism that practicable and helpful procedures will be developed.

Placement In Courses and Programs

At present, little attention is given to revising effective procedures for placing high school and college students in courses and programs where they can progress readily in their learning tasks without wasting time in duplicating what they have already learned or failing because they lack some necessary prerequisites for doing successful work. In a few high schools there are placement programs in English and foreign languages. A larger number of colleges have placement procedures in freshman English, mathematics, and foreign languages. There is much wider recognition of the potential value of placement programs. Part of the discrepancy between belief and practice is due to the lack of imaginative provisions for students whose previous learning permits advanced placement; but a major factor accounting for the limited number of programs is the difficulty of assessing the extent of a student's relevant previous learning. He may be prepared to tackle more advanced work in one or more of three respects: (1) He may have acquired and can use some of the modes of intellectual inquiry which are involved in the subject—for example, in literature he may know how to read and interpret novels, or in biology, he may know how to set up and perform experiments in the laboratory. (2) He may have developed an understanding of some of the basic concepts, generalizations, facts, or ideas which are involved in the subject. (3) He may be able to acquire the prerequisites to advanced work very quickly although he cannot demonstrate competence in them at the beginning. Some students have learned them earlier but have since forgotten. Others may have learned relevant things which can be "transferred" with further understanding and practice.

Our current placement tests are planned to sample content and sometimes to involve appropriate modes of inquiry. Forced by the need to economize on time in testing, scoring, and summarizing results, these tests have been too short to provide separate reliable scores for modes of inquiry and for content. Furthermore, they have usually been planned with the idea that a test approximating the final examination in the first course would provide scores which would indicate which students at entrance to the course could pass the final test. They are placed in more advanced courses.

Such a procedure works best when the student has actually taken the very same course earlier as is the case with the Advanced Placement Program operated jointly by certain schools and colleges. Such a test does not provide a dependable identification of students who have acquired the appropriate modes of inquiry but do not know the content used in the course, nor those who have developed an understanding of some of the basic concepts, generalizations, and the like. The procedure does not sort out the students who can quickly acquire or re-acquire the prerequisites to more advanced courses or programs.

The new machines furnish a practicable means for analyzing and summarizing test data which would serve these three purposes. The placement test can be built to sample systematically the student's ability to carry on the modes of inquiry involved in the subject, also to sample systematically the student's understanding of basic concepts, generalizations, facts, ideas, and the like. Furthermore, in the first few days of instruction, learning tasks can be assigned which can serve as a test of ability to re-learn or quickly acquire prerequisite skills and understanding. The computer can handle individual items as well as sub-scores and make a variety of transformations so as to estimate what skills and understandings the student has already acquired and to predict what he can now efficiently learn. These predictions may point to the need for several different class sections or learning programs to provide efficient placement, but such a variety of opportunities is feasible in large institutions.

Diagnosing Learning Difficulties

The foregoing comments on placement programs are also largely relevant to the diagnosis of learning difficulties since they are two closely related processes. In placement, we are estimating the learning needs of the student and the learning tasks which he can carry out efficiently. In diagnosis, we are concerned with students who are not learning effectively and efficiently and we seek to identify factors that will help to suggest learning tasks which they can carry on. In the latter case, we need to do the following:

(1) Identify students who are not learning effectively and

efficiently. Periodic tests or critical review of the students' learning exercises will normally serve this purpose. Programmed materials are perhaps easiest to check. Objective tests take little more time, and even the review of written essays simply to identify students who are not learning effectively is not very laborious. (2) Identify the learning exercises that are not being done efficiently, and seek to categorize them in ways which will help to locate the learning difficulty. (3) Use further testing to verify the hypotheses generated by (2).

It is in steps (2) and (3) that the new machines can be very helpful. The student's performance on every learning exercise can be recorded and a variety of categories can be employed in classifying successes and failures on each of them. For example, the simplest kind of categories would put together errors involving the same topic, and errors involving the same kind of intellectual behavior. If, by helping the student to understand which topics or which abilities he was having trouble with, further checking showed that he no longer made these errors, the diagnosis would be adequate. In some cases, however, a classification of this sort does not directly indicate the root of the difficulty the student is encountering. For example, he may have certain habits which interfere or certain emotional blocks, or an over-rigid conception of the subject. Additional tests may be necessary to reveal these less obvious factors. The main point here is that data of all these types can be used as input to the machines. Then, as experience accumulates, equations can be developed for estimating the probable consequences of each of several types of remediation which might be suggested by a specific diagnosis. This kind of information can help in building up practical diagnostic and remedial procedures in large schools where they are now rarely found.

Conclusion

These four kinds of educational problems — allocating and scheduling learning resources, admitting students to college, placing them in courses and programs, and diagnosing learning difficulties — are only illustrative of the areas in which the new machines can play a helpful role in educational decision-making. Their value largely lies in the speed and economy with which they can handle large bodies of information to furnish multivariable predictions of probable consequences of courses of action. These machines are expensive and they must be employed on large masses of data to obtain low unit costs. Large schools and colleges may be able to provide enough work to keep a high speed computer busy. For most schools and colleges, however, the values of the machines can be obtained economically only by joining with others in a group service. Fortunately, such service organizations are being formed so that more and more schools and colleges will be able to benefit from them at a reasonable cost. If intelligent programs are developed and followed, an important contribution will be made to American education.

Participants — 1962 Invitational Conference on Testing Problems

ABRAHAM, Ansley A., Florida Agricultural and Mechanical University
ADRIAN, Robert J., St. John's University, Jamaica, New York
AHRENS, Dolores F., Educational Testing Service
ALBRIGHT, Frank S., West Orange (New Jersey) Public Schools
ALLISON, Roger B., Jr., University of Utah Medical School
ALT, Pauline M., Central Connecticut State College
ALTENHAUS, Corrinne B., Orange (New Jersey) Public Schools
AMES, John L., Queens College
ANASTASI, Anne, Fordham University
ANDERSON, G. Ernest, Jr., Newton Public Schools, Newtonville, Massachusetts
ANDERSON, Rose G., New York City
ANDERSON, Roy N., North Carolina State College
ANDERSON, Scarvia B., Educational Testing Service
ANDREE, Robert G., Rich Township High School, Park Forest, Illinois
ANGOFF, William H., Educational Testing Service
ANNE, Sister Elizabeth, C.S.N., Valhalla, New York
ANTHONY, Brother E., F.S.C., La Salle College
AQUINAS, Sister Thomas, Seattle University
ARCHER, N. Sidney, Pennsylvania Department of Public Instruction
ARSENIAN, Seth, Springfield College, Massachusetts
ASH, Philip, Inland Steel, Chicago, Illinois
ASHLEY, Franklyn, Vernon Court Junior College, Newport, Rhode Island
ASTEN, Mildred, New York University
ATKINS, William H., Yeshiva University
ATWATER, Stephen K., Knox College

BALLISTY, Aloysius, Philadelphia Personnel Department
BANNON, Charles J., Department of Education, Waterbury, Connecticut
BANNON, Marion K., Department of Education, Waterbury, Connecticut
BARDACK, Herbert D., New York State Department of Social Welfare
BARGMANN, Rolf E., International Business Machines Research
BARKER, Roland, Montclair (New Jersey) Public Schools
BARNETTE, Edmund L., South Carolina State Department of Education
BARNETTE, W. L., Jr., University of Buffalo
BARRETT, Dorothy M., Hunter College
BARRON, Arleen S., Educational Testing Service
BARTNIK, Robert, Educational Testing Service
BATES, Margaret A., Educational Testing Service
BENNETT, George K., The Psychological Corporation
BENNETT, Marjorie G., Bronxville, New York
BENSON, Arthur L., Educational Testing Service

BENTZ, Jon, Sears, Roebuck & Co., Chicago, Illinois
BERG, Harry D., Michigan State University
BERG, Joel, King Philip School, West Hartford, Connecticut
BERGER, Bernard, New York City Department of Personnel
BERGESEN, B. E., Jr., Personnel Press, Inc., Princeton, New Jersey
BERGSTEIN, Harry B., Simpson Junior High School, Huntington, L. I., New York
BERKOWITZ, Edward, Audio-Visual Development and Appraisal Center, New York City
BEST, Phillip J., Educational Testing Service
BICKFORD, Lydia, Associated Press
BIGGERS, Julian, Texas Education Agency
BIRNEY, Robert C., Amherst College
BLACK, Donald B., University of Alberta
BLACK, Hillel, Jackson Heights, New York
BLEEKE, Donald E., Cranford (New Jersey) High School
BLIGH, Harold F., Harcourt, Brace and World, Inc.
BLOCK, Elaine C., Hofstra College
BLOCK, Renee, New York University
BLOOM, B. S., University of Chicago
BLUM, Stuart H., Hofstra College
BOBBITT, W. L., Charlotte-Mecklenburg (North Carolina) Schools
BOGART, Max, New Jersey State Department of Education
BOLAND, Ruth F., Cambridge (Massachusetts) School Department
BOLLENBACHER, Joan, Cincinnati (Ohio) Public Schools
BOROWITZ, Ruth, New York University
BORROWMAN, Ellen, Psychological Abstracts
BORROWMAN, Merle, Study of Teacher Education, New York
BOWEN, Mary, Educational Testing Service
BOYD, J. L., Educational Testing Service
BOYLSTON, Marie, Portland State College
BRACA, Susan E., Oceanside (New York) Senior High School
BRANDT, Hyman, New York City
BRANSFORD, T. L., Civil Service Department, New York
BRAUND, Robert A., Educational Testing Service
BRETNALL, William B., Educational Testing Service
BRISTOW, William H., Bureau of Curriculum Research, New York City
BROBST, Harry K., Oklahoma State University
BRODY, Sylvia B., New York University
BROSNAHAN, Sister M. Jonathan, College of St. Theresa
BROWN, F. Martin, Fountain Valley School, Colorado Springs, Colorado
BROWN, Frederick S., Great Neck (New York) Public Schools
BROWNE, Michael, Educational Testing Service
BRUDER, Gail Ann, Educational Testing Service
BRYAN, J. Ned, United States Office of Education
BRYAN, Miriam M., Educational Testing Service
BUDDEKE, Sister Rita, Catholic University

BUHLER, Roald, Educational Testing Service
BUKBERG, Alice, Plainview (New York) High School
BURKE, James M., Norwalk (Connecticut) Public Schools
BURKE, P. J., Bell Telephone Laboratories, Inc.
BURNHAM, Paul S., Yale University
BURNS, Hobert W., Syracuse University
BUROS, Oscar K., Rutgers, The State University, New Jersey
BURR, W. L., Harcourt, Brace and World, Inc.
BURTON, R. L., University of Oklahoma
BUSACKER, William E., Herkimer County Board of Cooperative Educational Services, New York
BUSSIS, Anne, Educational Testing Service
BUXENBAUM, Harris, New York University

CAFFREY, John, Harvard University
CALECHMAN, Pearl, New York University
CAMPBELL, Donald W., Newark (New Jersey) Public Schools
CANNON, Deore J., College of William and Mary
CAPPS, Marian P., Virginia State College
CARDINET, Jean, University of Illinois
CARNEGIE, Elizabeth, Nursing Outlook, New York City
CARSTATER, Eugene D., Bureau of Naval Personnel
CARSTATER, Marie H., Falls Church (Virginia) High School
CASPER, Thomas P., Louisville Catholic School Board, Kentucky
CHAMBERS, William M., University of Kentucky
CHANDLER, Thomas E., USA Southeastern Signal School, Fort Gordon, Georgia
CHAPPELL, Bartlett E. S., New York Military Academy
CHASE, C. I., Indiana University
CHAUNCEY, Henry, Educational Testing Service
CHAUSOW, Hymen M., TV College
CHEERS, Arlynne L., Grambling College
CHESNUT, Cora Mae, Rochester (New York) City School District
CHICKERING, Arthur W., Goddard College
CHILDS, E. Kitch, University of Chicago
CHUNG, Bom Mo, University of Chicago
CLARK, Philip I., California Test Bureau, Scarsdale, New York
CLEARY, Robert, Webster College
CLENDENEN, Dorothy M., The Psychological Corporation
COHEN, Nancy K., Pratt Institute
COLLEEN, Sister Mary, S.S.N.D., Archdiocese of New Orleans (Louisiana)
COLSON, John, Educational Testing Service
COMBS, Charles F., Board of Cooperative Educational Services, New York
CONLON, John, Northern Valley Regional High at Demarest, New Jersey

CONNELL, Ellen F., Erie (Pennsylvania) School District
CONNOLLY, John A., Educational Testing Service
CONSTANCE, Sister Mary, St. Vincent's Hospital, New York City
COPELAND, Herman A., State Civil Service Commission, Pennsylvania
CORY, Bertha Harper, United States Army Personnel Research Office
CORY, Charles H., Philadelphia Personnel Department
COSTA, Rena, New York University
COULSON, R. W., Butler University
COX, Henry M., The University of Nebraska
CRAVEN, Ethel Case, Veterans Administration, Brooklyn, New York
CREAGER, John A., National Academy of Sciences
CROOK, Frances, Schenectady, New York
CROSSON, Wilhelmina M., Palmer Memorial Institute
CURETON, Edward E., University of Tennessee
CURETON, Louise W., Knoxville, Tennessee
CURRY, John C., Mountain High School, West Orange, New Jersey
CURRY, Robert P., Cincinnati Board of Education
CURTIN, James R., United States Navy Research Activity
CURTIN, Wylma R., The Catholic University of America
CUTLIFF, Lucy Reed, Educational Testing Service
CUTTITTA, Frederick, New York City Board of Education
CYNAMON, Manuel, Brooklyn College

DAILEY, John T., American Institute for Research
DALLOW, Bob, Hauppauge (New York) High School
DALY, Alice, New York State Education Department
DAMARIN, Fred L., Jr., Educational Testing Service
DANA, Richard H., West Virginia University
DAVE, R. H., University of Chicago
DAVIDOFF, Melvin D., United States Civil Service Commission
DAVIDSON, Helen H., The City College of New York
DAVIS, Jay, Educational Testing Service
DAWSON, R. I., Equitable Life Assurance Society
DEAN, E. Douglas, Newark College of Engineering
DE BRULER, Fred M., Educational Testing Service, Seattle
DE CARLO, Charles R., International Business Machines Corporation
DEES, Bowen C., National Science Foundation
DETCHEN, Lily, Chatham College
DEVLIN, Nancy, Educational Testing Service
DIAMOND, Esther E., Science Research Associates
DIEDERICH, Paul B., Educational Testing Service
DIEDRICH, William A., Ulster County Guidance Association, New York
DIEHL, Mary Jane, Educational Testing Service
DIERS, Helen A., Long Island University
DION, Robert, California Test Bureau, Nashua, New Hampshire
DIZNEY, Henry, Kent State University

DODD, William E., International Business Machines Corporation
DONLON, Thomas F., Educational Testing Service
DOPPELT, Jerome E., The Psychological Corporation
DOWNES, Margaret, New York State Education Department
DRAGOSITZ, Anna, Educational Testing Service
DREW, Mary L., Educational Testing Service
DUBNICK, Lester, Huntington (Long Island, New York) Public Schools
DUDEK, Edmund E., United States Navy Personnel Research Activity
DUGAN, Robert D., American Institute for Research
DUNCANSON, James P., Educational Testing Service
DUNN, Frances E., Brown University
DUNNE, Reverend William J., S.J., National Catholic Educational Association
DUROST, Walter N., Test Service and Advisement Center
DUTTON, Eugene, University of Alabama
DWYER, Mary I., Pennsbury (Pennsylvania) Senior High School
DYER, Henry, Educational Testing Service

EBEL, Robert L., Educational Testing Service
EGBERT, Robert L., System Development Corporation
EHRMANN, Irene E., Verona (New Jersey) High School
EICHLER, Herbert, Columbia University
EKSTROM, Ruth B., Educational Testing Service
ELAINE, Sister Mary., S.S.N.D., Mt. Carmel (Texas) High School
ELLER, William, University of Buffalo
ELLIOTT, John M., The Richardson Foundation
ENGELHART, Max D., Chicago City Junior College
EPSTEIN, Bertram, City College of New York
EPSTEIN, Marion, Educational Testing Service
EVERETT, Margaret P., Educational Testing Service

FAGIN, Barry, Marquette University
FAN, Chung T., International Business Machines Corporation
FARESE, Joseph, Pennsbury (Pennsylvania) High School
FARR, S. David, University of Buffalo
FAY, Paul J., New York State Civil Service Department
FEINMAN, Herb, New York University
FELDMANN, Shirley, New York Medical College
FELDT, Leonard S., State University of Iowa
FENDRICK, David, Educational Testing Service
FENDRICK, Paul, Western Electric Company
FENOLLOSA, George M., Houghton Mifflin Company
FENSTERMACHER, Guy M., Educational Testing Service
FERGUSON, John L., University of Missouri
FERGUSON, Leonard W., Life Insurance Agency Management Association
FIELDER, Earl R., Educational Testing Service

FIFER, Gordon, Hunter College
FINDLEY, Warren G., University of Georgia
FINEGAN, Owen T., Gannon College
FLANAGAN, John C., American Institute for Research
FLANAGAN, Mrs. John C., Pittsburgh, Pennsylvania
FLEISCH, Sylvia, Boston University Computing Center
FLESCH, Mary H., Educational Testing Service
FLOYD, William A., Columbia College
FORMICA, Louis, Norwalk (Connecticut) Schools
FOX, David, Columbia University
FOXE, Mrs. Esther K., University of Maryland
FRANKFELDT, Eli, United States Army Personnel Research Office
FRENCH, John W., Educational Testing Service
FRENCH, Robert L., Science Research Associates, Inc.
FRICKE, Benno G., University of Michigan
FRIEDENBERG, Edgar Z., Brooklyn College
FRIEDMAN, Jerry J., Valley Regional High School, Deep River, Connecticut
FRIEDMAN, Sidney, Bureau of Naval Personnel
FRITZKY, Ferdinand, Educational Testing Service
FROEHLICH, G. J., University of Illinois
FUCHS, Morton, Jamaica High School, New York

GALLAGHER, H. L., Educational Testing Service
GAMES, Paul A., Ohio University
GARDNER, Eric F., Syracuse University
GARNER, William C., Dean Junior College
GEE, Helen H., University of Oregon
GERJUOY, Herbert, Educational Testing Service
GERBERICH, J. Raymond, American Educational Research Association
GIBBS, Cecil A., Educational Testing Service
GIBLETTE, John F., University of Maryland
GIDEON, Sara B., Pennsylvania State Department of Public Instruction
GLASER, Robert, University of Pittsburgh and American Institute for Research
GLASIER, Charles A., New York State Education Department
GLEASON, Walter P., Weston (Massachusetts) Public Schools
GOBETZ, Wallace, New York University Testing and Advisement Center
GODSHALK, Fred, Educational Testing Service
GOLD, Edward L., New York University
GOLDBERG, Miriam L., Columbia University
GOLDMAN, Leo, Brooklyn College
GOLDSCHMIDT, John E., Herricks Senior High School, Long Island, New York
GOLDSTEIN, Leo S., Center for Programed Instruction, New York City
GOLDSTEIN, William, Plainview Junior High School, Long Island, New York

GOOLER, Edward, Bay Shore (New York) Schools
GORDON, Leonard V., United States Army Personnel Research Office
GOREN, Arnold L., New York University
GORHAM, William, United States Civil Service Commission
GOSLIN, David A., Russell Sage Foundation, New York
GOTKIN, Lassar G., Center for Programed Instruction, New York City
GOULD, Kenneth M., Scholastic Magazines, New York City
GOWING, James D., Tabor Academy, Marion, Massachusetts
GRAFF, Franklyn A., Westport (Connecticut) Public Schools
GRAHAM, John, Educational Testing Service
GRANT, Donald L., American Telephone and Telegraph Company
GRAY, Lyle Blaine, Remedial Reading Consultant, Baltimore, Maryland
GREEN, Dorothy, United States Civil Service Commission
GREEN, Russel F., University of Rochester
GRIBBONS, Warren, Regis College
GROSE, Robert F., Amherst College
GROVES, Kenneth J., Air University, Maxwell Air Force Base, Alabama
GUILFORD, J. P., University of Southern California
GULLIKSEN, Harold, Educational Testing Service

HAAGEN, C. Hess, Wesleyan University
HAGEN, Elizabeth, Columbia University
HAGGERTY, Helen R., Army Personnel Research Office
HAGMAN, Elmer R., Greenwich (Connecticut) Public Schools
HALL, Alfred, Educational Testing Service
HALL, Robert, Manter Hall School, Cambridge, Massachusetts
HARMON, Lindsey R., National Academy of Sciences
HAROOTUNIAN, Berj, University of Delaware
HART, Josephine, New York State Civil Service Department
HARTSHORNE, Nathaniel H., Educational Testing Service
HARVEY, Philip R., Educational Testing Service
HAUBRICH, Vernon F., Hunter College
HAUSMAN, Howard J., National Science Foundation
HAWKINS, Connie, Ginn and Company
HAYWARD, Priscilla, Educational Testing Service
HEALY, John, North Junior High School, Bloomfield, New Jersey
HEATH, Robert W., Educational Testing Service
HEATON, Kenneth L., Heaton, Floyd and Watson, Philadelphia
HEFLIN, Virginia B., The City College of New York
HEIDGERD, Lloyd, Educational Testing Service
HEIDIG, Reverend Francis R., Archdiocese of Washington (D. C.)
HEISER, Ruth, Consulting Psychologist, Cincinnati, Ohio
HELMER, Etta D., Brooklyn College
HEMPHILL, John K., Educational Testing Service
HENRY, Sallyann, The Psychological Corporation
HENRYSSON, Sten, Bureau of Educational Research
HERGE, Henry C., Rutgers, The State University, New Jersey

HERMAN, David O., The Psychological Corporation
HERRICK, C. James, Hartwick College
HERZOG, Allen N., University of Chicago
HESLIN, Phyllis, National League for Nursing
HIERONYMUS, A. N., University of Iowa
HIGGINS, J. C., Fordham University
HILL, Walker H., Michigan State University
HILLS, John R., University System of Georgia
HILTON, Thomas L., Educational Testing Service
HITT, Fred J., New York State Education Department
HOBERMAN, Solomon, New York City Department of Personnel
HOFFMAN, E. Lee, Tulane University
HOLLIS, Esther R., The Psychological Corporation
HOLLISTER, John S., Educational Testing Service
HONAKER, Linton R., Tuscarawas County Schools, Ohio
HOOD, George W., Stetson University
HOPKINS, Florence M., Educational Testing Service
HOPMANN, Robert P., The Lutheran Church – Missouri Synod
HOROWITZ, Milton W., Queens College
HORTON, Clark W., Dartmouth College
HOWARD, Allen H., University of Illinois
HOWELL, John J., Educational Testing Service
HUDDLESTON, Edith M., United States Office of Education
HUDEN, John, University of Vermont
HUMPHRY, Betty J., Educational Testing Service
HUNT, Donald R., Suffolk County Civil Service Commission, New York
HUTCHINSON, Mary Gordon, Wilby High School, Waterbury, Connecticut
HUYSER, Robert, Educational Testing Service
HUYSER, Sarah L., Educational Testing Service
HYPPS, Irene C., Public Schools of the District of Columbia

IMPELLIZZERI, Irene H., Brooklyn College
IRBY, Alice J., Educational Testing Service

JACKA, Elizabeth, Illinois State Scholarship Commission
JACKSON, George S., Westbrook Junior College
JACKSON, Robert W. B., Department of Educational Research, Toronto, Canada
JACOBS, Paul, Educational Testing Service
JACOBS, Robert, Southern Illinois University
JACOBSON, Audrey K., W. E. Howitt Junior High School, Long Island, New York
JAMES, Grace R., University of North Carolina
JANEBA, Hugo B., Rutherford (New Jersey) High School
JARMUS, Daniel, New York University
JASPEN, Nathan, New York University

JENNINGS, Frank, Saturday Review Education Supplement, New York City
JOEL, D. E., St. Francis College
JOHNSON, A. Pemberton, Newark College of Engineering
JOHNSON, Lewis W., City of Philadelphia Personnel Department
JOHNSON, Victor B., State Department of Education, Florida
JORDAN, Arthur M., University of North Carolina
JUOLA, Arvo E., Michigan State University

KABACK, Goldie Ruth, The City College of New York
KALIKOW, Theodora J., St. Francis College
KAMMAN, James F., University of Illinois
KARAS, Shawky, Educational Testing Service
KARL, Madeline, New York City Board of Education
KATZ, Martin, Educational Testing Service
KATZELL, Mrs. Raymond A., National League for Nursing
KAUFFMAN, Eleonora J., University of Chicago
KAVRUCK, Samuel, United States Office of Education
KELEHER, Reverend Gregory, St. Anselm's College
KELLER, Franklin J., New York City
KELLEY, H. Paul, University of Texas
KENDALL, W. E., The Psychological Corporation
KENOSIAN, Joyce McCollum, New York State Education Department
KENT, Clarence L., State Department of Education, Virginia
KERSTING, Ethel, Educational Testing Service
KING, Richard G., Harvard College
KIP, R. de R., American Institute for Property and Liability Underwriters, Inc.
KLANIT, Frederick, New York City Community College
KLINE, William E., Baltimore County Public Schools
KLING, Frederick R., Educational Testing Service
KNAG, Howard A., Queens College
KNAPP, Thomas R., University of Rochester
KOCH, John C., Jr., Madison (New Jersey) High School
KOGAN, Leonard S., Community Service Society, New York City
KOGAN, Nathan, Educational Testing Service
KOOB, Reverend C. Albert, National Catholic Educational Association
KOPPE, William A., Character Research Project
KORNREICH, Jerome S., Microwave Associates, Inc.
KOSTUK, Max, State College at Boston
KOWITZ, Gerald T., New York State Education Department
KRIGSMAN, Ruben, New York City Community College
KUNOFSKY, Norma, Civil Service Department, New York
KUNOFSKY, Solomon, New York State Department of Health
KURFMAN, Dana, Educational Testing Service
KURLAND, Norman, New York State Education Department

LAKE, Barbara W., Fordham University
LALIBERTE, Richard A., Hinsdale (Illinois) Schools
LAMBERT, Jane E., Educational Testing Service
LANE, Hugh W., University of Chicago
LANGBERG, George, Hunter College
LANGMUIR, C. R., The Psychological Corporation
LANKTON, Robert S., Detroit Public Schools
LANNHOLM, Gerald V., Educational Testing Service
LAPETINA, Frank, Peekskill (New York) High School
LARSON, Cedric A., Rutgers, The State University, New Jersey
LAUDAN, Bonnie S., Educational Testing Service
LAWRENCE, Sister Helen, Trinity College
LEIDNER, Burton R., International Business Machines Corporation
LEIMAN, John M., System Development Corporation
LENNON, Roger T., Harcourt, Brace and World, Inc.
LEVINE, Harold G., New York State Education Department
LEVINE, Richard, Educational Testing Service
LIDA, S. L., International Business Machines Corporation
LIGON, Ernest M., Union College
LINDEMAN, Richard H., Columbia University
LINDSEY, John R., Detroit (Michigan) Public Schools
LINK, Frances R., Cheltenham Township (Pennsylvania) Public Schools
LITTERICK, William S., Keuka College
LOHNES, Paul R., University of New Hampshire
LONG, Lillian D., Professional Examination Service, New York
LONG, Louis, City College of New York
LORD, Frederic M., Educational Testing Service
LORET, Peter G., Educational Testing Service
LORGE, Sarah W., Language Laboratory Project, New York
LOTTES, John, State University College, New York
LOWERY, Zeb A., College Board Test Center, Rutherfordton, North Carolina
LUCAS, Diana D., Educational Testing Service
LUSK, Louis T., Independent Vocational Counselor, Norwalk, Connecticut
LYMAN, Howard B., University of Cincinnati

MacBAIN, Robert T., The Torrington Company, Connecticut
MADDEN, Richard, Sonoma State College
MADDOX, Clifford R., Cedarville College
MADISON, Peter, Princeton University
MAHAN, Thomas W., Jr., University of Connecticut
MAIER, Milton, Educational Testing Service
MALCOLM, Donald, Educational Testing Service
MALINCHOC, Joseph, Educational Testing Service
MALONEY, Daniel J., New York State Education Department
MALTBY, Jane M., Yale University

MANLEY, Myron, Educational Testing Service
MANNELLO, George, Jr., Hofstra College
MAROHN, Mildred, Lutheran Church, St. Louis, Missouri
MARRON, Joseph E., United States Military Academy
MARSH, James, Educational Testing Service
MARSHALL, Randall, State Department of Education, New Jersey
MASLOW, Albert P., United States Civil Service Commission
MASON, Andrew W., Maryland State Department of Education
MASON, Louise-Mary, The Pennsylvania State University
MATHEWS, Chester O., Ohio Wesleyan University
MATHEWSON, Robert, City University of New York
MATTHEWS, Frank M., Pennsylvania Military College
MAYERS, Edith C., Bureau of Audio-Visual Instruction, New York City
MAYO, Samuel T., Loyola University
McCALL, W. C., University of Southern California
McCANN, Forbes E., McCann Associates, Philadelphia
McCARTHY, Dorothea, Fordham University
McCLARY, George O., Richmond (Virginia) Public Schools
McCONNELL, John C., Windward School, White Plains, New York
McCONNELL, Thomas R., University of California
McCONVILLE, Carolyn B., Educational Testing Service
McCORD, Richard B., City of Philadelphia Personnel Department
McDILL, Mrs. Thomas, The Westminster Schools, Atlanta, Georgia
McINTIRE, Paul H., University of New Hampshire
McDANIEL, Sarah W., Charlotte-Mecklenburg (North Carolina) Schools
McKENNA, Mae R., Crosby High School, Waterbury, Connecticut
McKENZIE, F. W., Board of Education, Darien, Connecticut
McKEON, James J., Educational Testing Service
McLAIN, Joseph C., Mamaroneck, New York
McLAUGHLIN, Kathryn D., Educational Testing Service
McLAUGHLIN, Kenneth F., United States Office of Education
McQUEEN, Joan, Peekskill (New York) High School
MEADE, Martin, Fordham University
MEHTA, Prayag, National Institute of Education, New Delhi, India
MELIGI, El, New Jersey Neuro-Psychiatric Institute
MELVILLE, S. Donald, Educational Testing Service
MERWIN, Jack C., University of Minnesota
MIANO, Michael, Educational Research, Brooklyn, New York
MIDDENDORF, Lorna, Roosevelt Junior High School, Westfield, New Jersey
MILES, Mrs. John R., Falls Church (Virginia) High School
MILHOLLAND, John E., University of Michigan
MILLER, Howard G., North Carolina State College
MILLER, Marian B., State Department of Public Instruction, Delaware
MILLMAN, Jason, Cornell University
MIRKIN, Louise, Educational Testing Service

MITCHELL, Blythe C., Harcourt, Brace and World, Inc.
MITTMAN, Arthur, State University of Iowa
MOLDOVAN, Leonard, Junior High School 38, Bronx, New York
MOLLENKOPF, William G., Procter and Gamble Company
MOREAU, Reverend George H., O.M.I., Diocese of Buffalo, New York
MORIARTY, Doris, Educational Testing Service
MORRISON, Adair, General Electric Research Laboratory
MORRISON, Alexander W., Polytechnic Institute of Brooklyn
MORRISON, Donald R., Milwaukee City Service Commission
MOWRER, George E., St. Louis City Board of Education, Missouri
MULLIGAN, Alexander R., Plainview Public Schools, Farmingdale, New York
MULLIGAN, Mrs. A. R., Farmingdale, New York
MUNGER, A. M., The Psychological Corporation
MUNSON, S. E., Franklin and Marshall College
MYERS, Albert, Educational Testing Service
MYERS, Franklin G., Scarsdale (New York) High School
MYERS, Isabel Briggs, Swarthmore, Pennsylvania

NAFDE, G. H., National Institute of Education, New Delhi, India
NEEDELMAN, Bertram, City of Philadelphia Personnel Department
NELSON, Robert A., Rutherford County Board of Education, North Carolina
NEVIN, Margaret, Educational Testing Service
NOBLE, Janet, Scholastic Teacher Magazine, New York City
NOISEUX, Ethel R., Civil Service Department, New York
NOLAN, David M., Educational Testing Service
NOLL, Victor H., Michigan State University
NORA, Sister Mary, S.S.N.D., National Catholic Educational Association
NORRELL, Gwen, Michigan State University
NORTH, Robert D., Educational Records Bureau
NULTY, Francis X., Educational Testing Service

O'KEEFE, John J., Science Research Associates
OLDHAM, Francis H., Watchung Hills Regional High School, Plainfield, New Jersey
OLIVE, Helen T., Columbia University
ORR, David B., American Institute for Research
ORSHANSKY, Bernice, New Rochelle (New York) High School
OSTRAM, Elizabeth, Civil Service Department, New York
OTIS, C. Robert, California Test Bureau, New York
OTTOBRE, Frances M., Educational Testing Service
OVERALL, John E., Kansas State University

PACE, C. Robert, University of California
PACKARD, Albert G., Baltimore (Maryland) Public Schools
PALMER, Orville, Educational Testing Service
PARLOFF, M. B., National Institute of Mental Health

PASCIUTI, Edward C., New Rochelle (New York) High School
PAUFF, George W., Scarsdale (New York) High School
PENGEL, J. E., YMCA Counseling Service, Hartford, Connecticut
PENGEL, Marguerite H., YMCA Counseling Service, Hartford, Connecticut
PETERS, Frank R., State University of New York at Stonybrook
PETERSON, Donald A., Life Insurance Agency Management Association
PETERSON, Richard, Educational Testing Service
PFEIFFER, Ann, Baltimore (Maryland) Public Schools
PIDGEON, Douglas A., Columbia University
PITCHER, Barbara, Educational Testing Service
PITMAN, Sir James, London, England
PLUMLEE, Lynnette B., Sandia Corporation
POLLACK, Norman C., Civil Service Department, New York
PRESCOTT, George A., University of Maine
PRESTON, Braxton, Educational Testing Service
PUTNEY, Marjorie, Educational Testing Service

QUINN, John S., Jr., Harcourt, Brace and World, Inc.

RABINEAU, Louis, Pratt Institute
RAPPARLIE, John H., Owens-Illinois Glass Company, Toledo, Ohio
RAWAT, Daleep Singh, National Institute of Education, New Delhi, India
READ, Thomas, Hampton Roads Academy, Virginia
READING, Ethel, Educational Testing Service
READY, I. E., State Board of Education, North Carolina
REED, Bernard A., Trenton State College, New Jersey
REED, Marvin R., New Jersey Education Association
REID, Catherine F., Hunter College
REID, John W., Indiana State College
REILLY, Pamela C., American Institute for Research
REISS, Jean, Horton Watkins High School, St. Louis, Missouri
REUTER, William H., Educational Testing Service
REYNOLDS, George, Scarsdale (New York) Public Schools
RHUM, Gordon J., State College of Iowa
RICCI, Adriana, New York University
RICH, Manny, Educational Advance, New York
RICHARDINE, Sister Mary, B.V.M., National Catholic Educational Association
RICKS, James H., Jr., The Psychological Corporation
RIGHTHAND, Herbert, Connecticut State Department of Education
RIVERA-MEDINA, Eduardo J., University of Puerto Rico
ROBERTS, G. Tracey, State Civil Service Commission, Pennsylvania
ROBINSON, Placida, The Associated Negro Press
ROHRBAUGH, Frances, Educational Testing Service
ROMANELLA, Alan E., Port Jefferson (New York) Public Schools

ROSENBLOOM, Paul C., University of Minnesota
ROSENGARTEN, William, Jr., Roslyn (New York) Public Schools
ROSENTHAL, Elsa, Educational Testing Service
ROWDEN, Dorothy, Markle Foundation, New York City
RUBINFELD, William, Mountain High School, West Orange, New Jersey
RUMMEL, J. F., University of Oregon

SACHS, Lorraine P., National League for Nursing
SALZBERG, Florence H., East Memorial Elementary School, Farmingdale, New York
SANDERS, John H., Rice University
SANFORD, Ruth C., West Hempstead (New York) Junior-Senior High School
SANZOTTO, Samuel M., Department of Public Instruction, Pennsylvania
SASLOW, Max S., New York City Department of Personnel
SAUNDERS, D. R., Princeton, New Jersey
SAUPE, Joe L., Michigan State University
SCANLON, Edward C., College Entrance Examination Board
SCATES, Alice Y., United States Office of Education
SCHNEIDER, Harriet, National League for Nursing
SCHLEKAT, George, Educational Testing Service
SCHNITZEN, Joseph P., University of Houston
SCHOFIELD, Gordon T., Maumee Valley Country Day School, Ohio
SCHRADER, William B., Educational Testing Service
SCHULT, Henry, New York University
SCHULTZ, Charles, Educational Testing Service
SCHWARZ, Mariette, Educational Testing Service
SCOTT, C. Winfield, Rutgers, The State University, New Jersey
SCRIBNER, Peter C., Harcourt, Brace and World, Inc.
SEIBEL, Dean W., Educational Testing Service
SERAYDARIAN, Tom, Princeton (New Jersey) High School
SERLING, Albert M., Educational Testing Service
SETTE, William V., Plainfield (New Jersey) High School
SEYFRIT, Paul L., Department of Education, Kentucky
SFORZA, Richard F., United States Military Academy
SHAPIRO, Bernard J., Harvard University
SHARAR, Paul H., Vocational Service Center, New York City
SHARP, Catherine G., Educational Testing Service
SHAYCOFT, Marion F., American Institute for Research
SHEA, James E., State Civil Service Department, New York
SHEPARDSON, Marie E., Educational Testing Service
SHIELDS, Mary, National League for Nursing
SHIMBERG, Benjamin, Educational Testing Service
SIEGEL, M., New York University
SILVEY, H. M., State College of Iowa
SIMANDLE, Sidney, Kentucky State Department of Education

SIMON, Dorothy, New York City Department of Personnel
SIMPSON, Elizabeth A., Illinois Institute of Technology
SINGLETON, Carlton M., Paterson State College
SITGREAVES, Rosedith, Columbia University
SKAGER, Rodney, Educational Testing Service
SKIMMER, Charles A., Columbia University
SKINNER, James A., Corpus Christi (Texas) Public Schools
SMITH, Ann Z., Educational Testing Service
SMITH, Arthur E., National Merit Scholarship Corporation
SMITH, George A., Educational Testing Service
SMITH, Lyman, American Medical Association
SMITH, Marshall P., Trenton State College
SMITH, Marshall S., Harvard University
SMITH, Robert E., Educational Testing Service
SMOTHERS, James R., Jr., United Negro College Fund, Inc.
SNIFFEN, A. Mead, Board of Cooperative Educational Services, New York
SNYDER, Arthur, Jr., Weeks Junior High School, Newton Centre, Massachusetts
SOAR, Robert S., University of South Carolina
SOLOMON, Robert J., Educational Testing Service
SOMMER, John, Houghton Mifflin Company
SOUTHER, Mary Taylor, Tower Hill School, Wilmington, Delaware
SOUTHWORTH, J. Alfred, University of Massachusetts
SPAULDING, Geraldine, Educational Records Bureau
SPEER, George S., Illinois Institute of Technology
SPENCE, James R., State University of New York
SPENCE, Ralph B., Columbia University
SPENCER, Richard E., Pennsylvania State University
SPRAGUE, Arthur, Hunter College
SPREMULLI, Estelle, Educational Testing Service
STAHL, Suzanne, Educational Testing Service
STANLEY, Julian C., University of Wisconsin
STEFFEN, Hans, Prudential Insurance Company
STEINBERG, Joanna, The Personnel Laboratory, Incorporated
STEINMAN, Arthur M., Trenton State College
STERLING, G. Dykeman, Educational Testing Service
STEWART, Clifford T., University of South Florida
STEWART, E. Elizabeth, Educational Testing Service
STEWART, Naomi, Cranford, New Jersey
STICE, Glen, New York City
STICKELL, David W., Educational Testing Service
STILES, Grace Ellen, University of Rhode Island
STOKER, Howard W., Florida State University
STONE, Alice V., Bucks County Office of Education, Doylestown, Pennsylvania
STONE, Paul T., Huntingdon College

STORM, Jerome F., Hofstra College
STOUGHTON, Robert W., Connecticut State Department of Education
STRICKER, Lawrence J., Educational Testing Service
STUDDIFORD, Walter B., Princeton University
STUNKARD, Clayton L., New York State Education Department
SUBACH, Anthony, New York University
SUPER, Donald E., Columbia University
SUTTON, Joseph T., Stetson University
SWAN, Beverley B., Florida State Department of Education
SWANSON, Carl W., New Jersey State Department of Education
SWANSON, Edward O., University of Minnesota
SWINEFORD, Frances, Educational Testing Service

TABLER, Charles J., New Jersey Department of Education
TALIANA, Lawrence E., Southern Illinois University
TANZER, Charles, La Salle Junior High School, New York City
TAYLOR, Bruce, Educational Testing Service
TAYLOR, Justine, Educational Testing Service
TAYLOR, Marvin, Queens College
TAYLOR, Samuel J., New York State Personnel Department
TERRAL, J. E., Educational Testing Service
TERWILLIGER, James S., Educational Testing Service
THOMAS, William F., University of Wisconsin
THORNDIKE, Robert L., Columbia University
THORNE, Margaret A., Educational Testing Service
THURSTONE, Thelma G., University of North Carolina
TIEDEMAN, David V., Harvard University
TORREY, Frederick G., Mount Hermon (Massachusetts) School
TRAXLER, Arthur E., Educational Records Bureau
TREMER, Carl E., Civil Service Department, New York
TREXLER, Laura M., Northern Valley Regional High at Demarest, New Jersey
TRIGGS, Frances, Committee on Diagnostic Reading Tests, Inc.
TRUMP, Paul L., American College Testing Program
TUCKER, Ledyard R, University of Illinois
TULLY, G. Emerson, Florida State University
TURNBULL, William W., Educational Testing Service
TYLER, Gladys, Race Brook, Woodbridge, Connecticut
TYLER, Ralph W., Center for Advanced Study in the Behavioral Sciences

UMANS, Shelley, New York City Board of Education
URQUHART, Helen L., Brown University

VALLEY, John R., Educational Testing Service
VAN SCHOICK, Edith, New York University

VITELES, Morris S., University of Pennsylvania
VON FANGE, Erich A., Concordia College
VOYLES, Louis V., University of Florida

WAGNER, Paul, Pennsylvania State College
WAITE, Annette, Educational Testing Service
WALKER, R. N., Personnel Press, Princeton, New Jersey
WALLACE, Wimburn L., The Psychological Corporation
WALLER, Raymond L., Allentown School District, Pennsylvania
WALLMARK, Madeline, Educational Testing Service
WALSH, John J., Boston College
WALTHER, Regis, Department of State
WALTON, Arline J., Malden (Massachusetts) Public Schools
WALTON, Wesley W., Educational Testing Service
WALTZMAN, Hal, Civil Service Department, New York
WARD, Annie W., DeLand, Florida
WASSERMANN, Lorraine, Educational Testing Service
WATKINS, Richard, Educational Testing Service
WATSON, Walter S., The Cooper Union, New York City
WEBER, Burton J., Roosevelt Junior High School, Westfield, New Jersey
WEINER, Max, Brooklyn College
WEISBERG, Louis G., New York University
WEISS, Eleanor S., Educational Testing Service
WEITZ, Henry, Duke University
WELLCK, A. A., University of New Mexico
WESMAN, Alexander G., The Psychological Corporation
WESTEN, Risdon J., United States Air Force Academy
WEXLER, Amelia H., Hofstra College
WHALEY, F. L., Pennsylvania State University
WHARTON, William P., Allegheny College
WHITE, Herbert L., New York City Department of Personnel
WHITLA, Dean K., Harvard University
WHITNEY, King, Jr., The Personnel Laboratory
WICKER, Frank, Educational Testing Service
WILAND, Charles B., Tuscarawas County Schools, Dover, Ohio
WILKINS, Walter L., United States Navy Medical Neuropsychiatric Research Unit
WILLARD, Richard W., Massachusetts Institute of Technology
WILLEY, Clarence F., Norwich University
WILLIAMS, Malcolm J., United States Coast Guard Academy
WILLIAMS, Roger K., Morgan State College
WILLINGHAM, Warren W., Georgia Institute of Technology
WILLIS, William S., New York University
WILSON, Kenneth M., Southern Regional Education Board
WILSON, Marilyn, Educational Testing Service
WILSON, Mrs. Paul A., White Plains (New Jersey) High School
WINANS, David, New Jersey State Department of Education

WINGO, Alfred L., State Board of Education, Virginia
WINTERBOTTOM, John, Educational Testing Service
WISE, Harold L., Educational Testing Service
WOMER, Frank B., University of Michigan
WOOD, Dorothy A., University of North Carolina
WOOD, James D., City University of New York
WORTHINGTON, Ross, Overseas News Features and World Photographic Service
WRIGHT, Douglas C., Jr., Educational Testing Service
WRIGHT, Elizabeth S., Bucks County Office of Education, Doylestown, Pennsylvania
WRIGHT, Wilbur H., State University College at Geneseo, New York
WRIGHTSTONE, J. Wayne, New York City Board of Education

YONAMINE, Mac, New York University
YUNDT, John H., Moorestown (New Jersey) Public Schools

ZACCARIA, Lucy, University of Illinois
ZALKIND, Sheldon S., Baruch School, New York City
ZOLA, Eugene J., Niskayuna High School, Schenectady, New York
ZUCKER, Marilyn J., Hofstra College
ZUCKERMAN, Harold, New York City Board of Education